

Open Research Online

The Open University's repository of research publications and other research outputs

A Multidisciplinary Study Of Antecedents To Voluntary Knowledge Contribution Within Online Forums

Thesis

How to cite:

Huo, Qunying (2017). A Multidisciplinary Study Of Antecedents To Voluntary Knowledge Contribution Within Online Forums. PhD thesis The Open University.

For guidance on citations see [FAQs](#).

© 2017 The Author



<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Version: Version of Record

Link(s) to article on publisher's website:

<http://dx.doi.org/doi:10.21954/ou.ro.0000cdea>

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

oro.open.ac.uk

A multidisciplinary study of antecedents to voluntary knowledge contribution within online forums

A dissertation presented by

QunYing HUO

In partial fulfilment of the requirements for Degree of Doctoral of Philosophy

The Open University

Supervised by Dr. Adrian PALMER and Dr. Cláudia SIMÕES

Dec 31 2016

Table of contents

Table of contents	2
List of figure	10
List of table	13
Abstract	16
Acknowledgements	17
Author's declaration	18
Abbreviations	20
Glossary	22
Chapter 1 Introduction.....	27
1.1 Research context.....	27
1.2 Research aims and objectives	29
1.3 The present research.....	36
1.4 Principle contributions of thesis	43
1.5 Limitations	45
1. 6 Thesis outline	45
Chapter 2 Literature review.....	47
2.1 Introductory elements.....	48
2.1.1 Definition of online forums	48

2.1.2 Social dilemmas in online forums	52
2. 2 Rationale of using multimethod approach.....	54
2.3 Theoretical background	57
2.3.1 Conceptualizing online trust	57
2.3.1.1The levels of trust.....	58
2.3.1.2 The dimensionality of trust	59
2.3.2 Theory of critical mass and perceived critical pass.....	61
2.3.2.1 The importance of theory of critical mass	61
2.3.2.2 The importance of initial contributors in theory of critical mass.....	63
2.3.2.3 Perceived critical mass	64
2.3.3 Predicting continuous online knowledge sharing behaviour	66
2.3.3.1 Introduction of the theoretic background	66
2.3.3.2 Intention toward knowledge sharing within online forums.....	69
2.3.3.3 The determinants of intention	71
2.3.3.4 Contextual factors for online contribution.....	73
2.3.3.5 The causal relationships between antecedents	77
2. 3.4 The development of trust	80
2. 3.5 The evolution of perceive critical mass within online forums	82
2.3.5.1 Background	82

2.3.5.2 Barabasi and Albert's Model (BA model)	85
2.3.5.3 Theory of critical mass applied in scale-free networks	87
2.3.5.4 Phase transitions	88
2.3.5.5 Percolation theory.....	89
2.3.5.7 Critical point	92
2.3.5.8 The importance of critical mass members in sustaining online forums.....	94
2.4 Summary of chapter.....	96
Chapter 3 Methodology	98
3.1 Epistemological overview.....	99
3.2 Inductive, deductive and retroductive research approaches.....	101
3.3 Research design	103
3.4 Methodology of study one – online survey	106
3.4.1 Online survey design.....	106
3.4.1.1 Identification of constructs	106
3.4.1.2 Measurement scale development	107
3.4.1.3 Questionnaire form and sequence.....	114
3.4.1.4 Sampling procedure	115
3.4.2 Pilot test.....	119
3.4.3 Preliminary data analysis.....	121

3.4.3.1 Missing data and outliers	121
3.4.3.2 Descriptive analysis	122
3.4.3.3 Response rate and non-response analysis	123
3.4.3.4 One-way ANOVA	125
3.4.4 Structural equation modelling	126
3.4.4.1 Introduction to SEM	126
3.4.4.2 CB-SEM processing	128
3.4.4.3 Testing of model fit	137
3.4.4.4 CB-SEM estimation techniques approach	140
3.4.4.5 Mediation Analysis	141
3.5 Methodology of study two - evaluating online trust	146
3.5.1 Research aims	146
3.5.2 Data collection	147
3.5.3 Method	149
3.5.4 Data analysis	151
3.5.4.1 Coding	151
3.5.4.2 SVM	153
3.5.4.3 Using recurrent neural network in time series	155
3.6 Methodology of study three - the role of critical mass in sustaining online forums	159

3.6.1 Research aims	159
3.6.2 Data collection	160
3.6.3 Method	161
3.6.3.1 Exploring the structure of online forums	162
3.6.3.2 Identifying critical mass members	166
3.6.3.3 The critical point.....	167
3.6.4 Data visualisation software and analysis language	169

Chapter 4 Results and discussions: study one – examinations of causal factors associated with online voluntary contribution of knowledge.....	170
--	------------

4 .1 <i>Introduction</i>	170
4.2 <i>Pilot test</i>	170
4.3 <i>The main study</i>	171
4.3.1 Preliminary data analysis.....	172
4.3.1.1 Missing data and outliers	172
4.3.1.2 Data distribution	172
4.3.1.3 Descriptive data analysis.....	173
4.3.1.4 Non-responses analysis.....	177
4.3.1.5 Comparison of groups of respondents.....	181
4.3.2 Measurement validity.....	183

4.3.2.1 Exploratory factor analysis	184
4.3.2.2 Testing reliability	186
4.3.2.3 Confirmatory factor analysis	186
4.3.3 Developing the structural model	194
4.3.3.1 Model fit indices of structural model	195
4.3.3.2 Parameter estimates for the structural paths	196
4.3.3.3 Examining plausibility of models	198
4.3.4 Testing mediated effects and moderated effects	200
4.3.5 Summary of results	213
4.3.5.1 Discussion	214
4.3.5.2 Theoretical implications	217
4.3.5.3 Managerial considerations	219
4.3.5.4 Limitations and further research	220
Chapter 5 Results and discussions: Study two - the development of online trust	222
5.1 Development of themes	222
5.2 Exploratory analysis	225
5.3 Dimensionality of online institutional trust	227
5.4 Evolution of online trust	230
5.5 Conclusions	237

Chapter 6: Results and discussions of study three: the role of critical mass members in sustaining online forums	241
6.1 Introduction.....	241
6.2 Descriptive data	242
6.3 Exploration of “contribution” network structure	244
6.3.1 Online forum Stack Overflow is scale-free	244
6.3.2 Comparing with candidate distributions	245
6.3.3 Self-sustaining scale-free network and Theory of critical mass	248
6.4 Evaluating the role of critical mass members	250
6.5 Evolution of online forums	254
6.6 Conclusion	259
Chapter 7: Conclusions.....	263
7.1 Introduction.....	264
7.2 Empirical findings.....	266
7.3 Theoretical implications	273
7.4 Managerial considerations	277
7.5 Limitations of the study.....	280
7.6 Recommendations for future studies	281
7.7 Summary conclusions.....	282
Appendix 1: Questionnaire	284

Appendix 2: Examples of coding	296
Appendix 3: Posterior probabilities against paired predictors.....	302
Bibliography.....	303

List of figure

Figure 1: Rationale of this thesis	38
Figure 2: Three different perspectives of Web 2.0 concepts	50
Figure 3: Conceptual model	69
Figure 4: An example of the bell curve and the power law distribution and growth	85
Figure 5: Largest cluster and numbers of clusters	90
Figure 6: Largest cluster and number of clusters	91
Figure 7: An illustration of the Cayley tree	92
Figure 3.1: Research design: rationale of this thesis	98
Figure 8: Chronology of the research processes: mix data analyses integrated in mixed- methods	106
Figure 9: An example of year 2009: ability scaling	152
Figure 10: an example of year 2009: emerging themes	153
Figure 11: Recurrent neural networks for time series	158
Figure 12: the initial parallel model	191
Figure 13: The original structural model	194
Figure 14: Mediation model one: the mediation effects of perceived critical mass on subjective norms	202
Figure 15: Moderation model two: the moderation effects of perceived critical mass on subjective norms	204
Figure 16: Moderated mediation model three: trust in online forums effects path a	205

Figure 17: Moderated mediation model four: trust in online forums effects path b	206
Figure 18: Moderated mediation model five: trust in online forums effects on path a and path b	207
Figure 19: Moderated mediation model eight: Trust in members effects on path b	210
Figure 20: Final result	221
Figure 21: Data distribution and correlation	226
A: ability; B: benevolence; I: integrity; P: predictability	226
Figure 22: Components extracted	226
Figure 23: Fitted responses and MSE	227
Figure 24: Paired predictors	229
Figure 25: Evolution of overall trust in online forums	231
Figure 26: Goodness plot: evolution of overall online institutional trust	232
Figure 27: Linearity plot of overall trust in online forums	232
Figure 28: Evolution of ability	
Figure 29: Goodness plot: evolution of ability	234
Figure 30: Evolution of benevolence	234
Figure 31: Goodness plot: evolution of benevolence	235
Figure 32: Evolution of Integrity	235
Figure 33: Goodness plot: evolution of integrity	236
Figure 34: Evolution of predictability	236

Figure 35: Goodness plot: evolution of predictability	237
Figure 36: Descriptive data: an overview of edges and size growth	244
Figure 37: Degree distribution	245
Figure 38: Comparing with candidate degree distributions	246
Figure 39: Probability plot of network attacks	252
Figure 40: After removing 5% of critical mass members	254
Figure 41: The evolution of online forum	255
Figure 42: Probability density functions of the 4 time periods	256
Figure 43: Preferential attachment plot of the successive stages	257
Figure 44: Dispersion of degree plot	258
Figure 7.1: Revisiting the rationale of this thesis	263

List of table

Table 1: The dimensions of trust	59
Table 2: Determinants of TPB	68
Table 3: Degree distributions applied to networks	84
Table 4: Purposes of mixed-methods research	104
Table 5: constructs within the online survey	107
Table 6: Measurement of perceived critical mass	112
Table 7: Measurement of trust in online forums	113
Table 8: Measurement of trust in members	113
Table 9: An illustration of CB-SEM process	132
Table 10: Summarizes the conditions for construct validity	135
Table 11: Summary of the fit indices adopted in this study	140
Table 12: An example of scales used for rating the node “ability”	152
Table 13: Alpha test	171
Table 14: Sample profile	173
Table 15: Descriptive data of individual variable	174
Table 16 Descriptive data of constructs (in percentage %)	176
Table 17: Independent t-test: early and late respondents	178
Table 18: Independent t-test: comparing means between genders	182

Table 19: One-way ANOVA: comparing means between ages.....	182
Table 20: One-way ANOVA: comparing means between education levels.....	183
Table 21: indices KMO and Bartlett test.....	184
Table 22: Results of PCA	185
Table 23: discriminant validity.....	189
Table 24: model fit indices	192
Table 25: overall model fit indices.....	192
Table 26 : final measurement model.....	192
Table 28: Model fit indices of the hypothesized model	196
Table 29 Regression weights of the hypothesized model	197
Table 30 Regression weights of alternative model one	199
Table 31 Regression weights for alternative model two	199
Table 32 Comparing model fit indices	200
Table 33: Moderation/Mediation effects of perceived critical mass	207
Table 34: Model fit indices: model N° one~five	208
Table 35: Collinearity test.....	210
Table 36: Path estimations model N° six ~ eight	211
Table 37: Model fit indices: model N° six ~ eight	211
Table 38: Mediation/moderation tests on the relationship between TRM-TRC-ATT	212

Table 39: Model fit indices: model N° nine ~ eleven	213
Table 40: Numbers of comments by emerged themes, 2005--2010.....	224
Table 41: Comparing the kernel function	228
Table 42: Power law fit generated by Python	248
Table 43: Attending and random attack.....	251

Abstract

One challenge faced by online forums is the provision of a sustainable supply of contributions of knowledge (Wasco *et al.*, 2009). Previous studies have identified online trust and perceived critical mass as antecedents of online knowledge contributions. However, the dynamic aspects of antecedents are little investigated. Moreover, how the dynamics together impact on members' willingness to contribute knowledge is an open question to be further investigated.

To examine the dynamic antecedents of online knowledge continuance, this thesis seeks to develop a holistic approach through three studies. Drawing on a decomposed theory of planned behaviour (Taylor and Todd, 1995), study one identifies dynamic antecedents of intentional online contribution behaviours. Covariance-based structural equation modelling analysis of 910 responses obtained shows that perceived critical mass and trust in online forums that mediates trust in members are the highlighted antecedents in the context of online forums. The development of trust in online forums is investigated through a time series approach in study two. Findings using webnographic and machine learning analysis show that the cognitive dimension of institutional trust is essential in initial trust building. Study three uses network analysis techniques to explore the role of critical mass members. Results indicate that only 5% of critical mass members can sustain online forums. However, critical mass members compete for their connections, inferring the importance of brand building in the beginning of online forums development. A summary of findings from the three studies suggests that the structure assurance of online forums can mediate the effects of interactions between members to a coalition of membership over time. The study provides further knowledge on the voluntary contribution within online forums by taking a dynamic approach, while previous studies in this field are predominantly cross-sectional and un-prophetic.

Acknowledgements

I would like to express my utmost special gratitude to my supervisors for their excellent and insightful supervision. My supervisors, Dr. PALMER and Dr. SIMÕES have been extremely helpful in not only offering me the freedom to pursue my research interests but also maintaining rigorous standards. This is where I have benefited most in my research experiences at The Open University.

I would also like to thank Dr. Sarah HUDSON for her instrumental help and insightful feedbacks in my past annual report assessments. Meanwhile, my thanks leads to Dr. Hans BORGMAN for his understanding and support when coming to him for discussions related to my research during my final stage of my PhD thesis.

My great appreciations also go to Ms Helen CASTLEY for her capacity in the administrative communications. I truly thank Ms Su PRIOR and Ms Clair WYLDE for their great service as the doctoral programme administrators at The Open University, and to Ms Patricia FOUEL for her kind help as the administrative coordinator with The Open University.

I thank my third party monitors Dr. Graham Winch, Dr. Philip Kitchen and Dr. Owain SMOLOVIC-JONES for their encouragement and emotional support during my research experiences at The Open University.

Finally, this doctoral research is dedicated to my family in Chengdu, Sichuan province in China, whose love and support to me are second to none.

Author's declaration

I, QY HUO, declare that the ideas, research work, analyses and conclusions reported in my thesis for the Doctor of Philosophy titled “*A multidisciplinary study of antecedents to voluntary knowledge contribution within online forums*” are entirely my effort under the supervision of Dr. Adrian PALMER and Dr. Cláudia SIMÕES, except where otherwise acknowledged.

During the preparation of this thesis, a number of papers were prepared as listed below. The remaining parts of the thesis are unpublished.

Journal:

PALMER, A. and HUO, Q. (2013) A study of trust over time within a social network mediated environment. *Journal of Marketing Management*. 29(15-16). p. 1816-1833.

HUO, Q. and PALMER, A. (2015) Analysing voluntary contribution to online forums using a proposed critical mass contribution model. *Journal of Applied Business Research*. 31(2), 687-700.

Conference papers:

PALMER, A. and HUO Q. (2011) A Longitudinal analysis of trust in a fast growth technology-based company. *Proceedings of the 15th World Marketing Conferences*. Reims Management School, Reims, France. 19-21 July 2011.

PALMER, A. and HUO Q. (2012) A time series, webnographic analysis of trust in a viral high-tech brand. *Proceedings of the 12th International Research Seminar in Service Management*. Aix-Marseille Université, Aix-en-Provence, France, 28 May – 1 June 2012.

PALMER, A. and HUO Q. (2012) Critical-mass effects on contribution to online service fora. *Proceedings of the American Marketing Association SERVISIG conference*, Hanken University, Finland, 7 – 9 June 2012.

PALMER, A and. HUO Q. (2012) Critical mass as an antecedent to stability in online communities. *Proceedings of the International days of International Days of Statistics and Economics*. The University of Economics, Czech Republic. 10 – 12 September 2012.

HUO, Q. and PALMER, A. (2015) Modelling users' contribution of knowledge to online forums. *Proceedings of the 6th 6th International Research Meeting in Business and Management (IRMBAM-2015)*, IPAG Business School, Nice, France, 2-3 July 2015.

HUO, Q. and PALMER, A. (2016) An empirical study of online knowledge contribution continence: the social and structural influences. *Proceedings of the American Marketing Association SERVISIG conference*. Aix-Marseille Université, Aix-en-Provence, France, 28 May – 1 June 2016.

HUO, Q. and PALMER, A. (2016) Seeding targets within online forums. *Proceedings of World Marketing Conference*, IESEG Business School, Paris, France, 19 – 21 July 2016.

Abbreviations

CB-SEM: Covariance-based structural equation modelling. It is preferred when the sample size is big enough (e.g. Preacher *et al.*, 2007) and the latent constructs are multidimensional (e.g. Wright *et al.*, 2012).

CFA: confirmatory factor analysis measures whether the data can fit the hypothesized measurement model (e.g. Schumacker and Lomax, 2004).

CFI: comparative fit index assumes that all observed variables loading toward a factor are not correlated (Cooper *et al.*, 2008), and is another alternative to the chi-square tests (Hooper *et al.*, 2008).

DTPB: decomposed theory of planned behaviour proposed by Tylor and Todd (1995). DTPB decomposes the determinants identified in TPB in order to fit the different contexts (e.g. Tylor and Todd, 1995).

EFA: exploratory factor analysis is to maximize the variances shared among the observed indicators. In this thesis, it serves to identify the latent variables with a set of observed measures (e.g. Fabrigar *et al.*, 1999).

GFI: goodness of fit index is measured through the weighted sum of squared residuals from a covariance matrix, and the weighted sum of squared covariance and variance (Hoyle, 2014).

ML: maximum likelihood is defined as “a method of statistical estimation which seeks to identify the population parameters with a maximum likelihood of generating the observed sample distribution” (Lewis-Beck, 1994:153).

NNA: neural network analysis for time series is recognised as an unsupervised method for mapping the complex nonlinear relationships between the inputs X in order to predict the output Y in the following period (e.g. Nielson, 2016).

PCA: principle component analysis is associated with EFA in this thesis, and is the method to convert the possible correlated observed measures to linearly uncorrelated variables, namely principle components (e.g. Hair *et al.*, 2010).

PGFI: parsimony goodness of fit index is based on GFI by adjusting the loss of degree of freedom (Cooper *et al.*, 2008).

RMSEA: the root mean square error of approximation is an alternative measure to chi-square (χ^2) tests (Bentler, 2007).

SEM: the statistical analysis techniques that involve the confirmatory factor analysis and path analysis (e.g. Anderson and Gerbing, 1988).

SF: scale-free networks initially proposed by Barabasi and Albert (1999) are characterized by growth and preferential attachment.

SVM: support vector machine can map the nonlinear relationships between the predictors and response variable by projecting the multiple dimensional patterns into two half spaces (Caragea *et al.*, 2001).

TPB: theory of planned behaviour proposed by Ajzen (1991). Knowledge sharing is an intentional behaviour (e.g. Ajzen, 1991).

Glossary

Bias-correct percentile bootstrapping method: provides the statistical power to examine whether the indirect effects of mediation or moderations are significant or not (Preacher *et al.*, 2007). The bias is the difference between the expected testing error and training error (Steck and Jaakkola, 2003).

Chi-square tests: tests tend to measure whether the observed covariance matrix, Σ , is equal to the expected covariance matrix, $\Sigma(\theta)$ (Hoyle, 2014).

Collective actions: they involve voluntary contributions and are the solutions to the social dilemma (e.g. Ostrom 2000).

Common factors: factors that have influence on more than one observed indicators (e.g. Schumacker and Lomax, 2004).

Congeneric model: allows the true scores and variance for all items to be differ (e.g. Wright *et al.*, 2012).

Convergent validity: the measures of the degree to which the constructs that are theoretically hypothesized are related (e.g. Wright *et al.*, 2012).

Critical mass members: the exchange terms with hubs in this thesis.

Critical point: the point in which a mass phenomenon appears (e.g. Oliver and Marwell, 1988).

Cronbach's alpha (α): the commonly used technique within EFA for assessing internal consistency within measurement scales (e.g. Cortina, 1993).

Degree: the connections associated to nodes (Barabasi and Albert, 1999).

Direct network: the orientation of edges is considered within the direct networks (Ajith *et al.*, 2011).

Discriminant validity: the measures of the degree to which the constructs that are theoretically hypothesized are not related (e.g. Hair *et al.*, 2010).

Edges: an edge is created between the message poster and responder in this thesis.

Endogenous variables: the dependent latent variables Y that are measured through the coefficient matrix of the relationship between observed indicators toward the latent variables, and the associated error vector for endogenous variables (e.g. Schumacker and Lomax, 2004).

Exogenous variables: the independent variables X that are measured through the observed factor loadings to X , the common factors and the associated residual vector for exogenous variables (e.g. Schumacker and Lomax, 2004).

Face validity: associated with EFA in order to understand whether the latent constructs have both the theoretical and practical meanings (e.g. Hair *et al.*, 2010).

Factor rotation: interprets factors in multidimensional space to a simple structural factor matrix (e.g. Fabrigar *et al.*, 1999).

Hubs: nodes that hold the majority of connections within networks (e.g. Barabasi, 2013).

Indirect network: the orientation of edges is ignored within the indirect networks (Ajith *et al.*, 2011).

Jointness of supply: the cost of joint production of public goods is lower than would be the case if the goods were produced separately (Buchanan, 1966).

Latent variables: those variables those are not able to be directly observed but inferred from the directly observed variables (e.g. Borsboom *et al.*, 2003).

Mediation: refers to the causal effects of independent variable X on the dependant variable Y which are transmitted through the mediator M_e (Baron and Kenny, 1986).

Measurement models: the hypothesized models that examine the relationships between the endogenous variables Y and their measures, the exogenous variables X (e.g. Bryne, 2013).

Moderation: refers to the causal relationships between X and Y which are influenced by a third variable M_o , or the third variable M_o interacts with X regressing to Y (Baron and Kenny, 1986).

Moderated mediation: measures the mediation and moderation effects simultaneously, and is performed on continuous variables and for a big sample size (Preacher *et al.*, 2007).

Nodes: members within online forums in this thesis.

Non-excludability: all individuals in a collective having a right to access a public good, regardless of their contributions (Snidal, 1979).

Nonparametric bootstrapping: refers to randomly chosen samples taken from the initial data set for replacement to the bootstrapping data, and repeats this procedure of re-sampling k times (Preacher *et al.*, 2007).

Non-rivalry: goods that are not used up after consumption, such as the public park (Samuelson, 1954).

Oblique rotation: the factor rotation method that allows factors to be correlated, and is recommended when the correlation coefficients between paired factors are over than 0.32 (e.g. Fiddell, 2007).

Online forums: interest-oriented online communities based on the Web 2.0 technology which deals primarily with knowledge sharing by gathering members who share a common interest or theme (Spaulding, 2009).

Orthogonal rotations: the factor rotation method that assumes factors are not correlated (e.g. Gorsuch, 1983).

Parallel model: constrains that the true scores and error terms to all items are equal (e.g. Wright *et al.*, 2012).

Path analysis: the analysis of the regression relationships between latent constructs within a hypothesized model (e.g. Hair *et al.*, 2010).

Preferential attachment: the probability that the new node will connect an existing node proportionally depends on the degree of existed node (Barabasi and Albert, 1999).

Production function: the input often does not have an influence on the output of a public good (e.g. Ostrom, 2010). A u-concave or accelerating function ensures the increasing marginal returns.

Public goods: the four dimensions that characterize public goods are non-rivalry, non-excludability, the production function and jointness of supply (Wasco *et al.*, 2009).

Reliability: the measures of the degree to which an assessment can produce stable and consistence results (e.g. Hair *et al.*, 2010).

Social dilemma: occurs when the public good has both non-rivalry and non-excludability (Ostrom, 2000; Ostrom, 2010).

Spinning cluster: associated with the infinite system (e.g. Newman, 2005), and almost immediately appears after the critical point through which nodes within the network are connected (e.g. Cohen *et al.*, 2002).

Structure of online forums: measured through the degree distribution of the online forum under study in this thesis.

Sustainability: the goal of the sustainable development. It is often seen as a holistic approach involving aggregated factors that enable sustainable development (e.g. Sheth *et al.*, 2010).

Sustainable development: the dynamic process influenced by the numbers of identified factors to achieve the goal of “sustainability” (e.g. Diesendorf, 2000).

Sustainability of online forums: the aggregated factors that enable the knowledge continuance provided within online forums over time.

Tau model: constrains that the true scores for all items are equal (e.g. Wright *et al.*, 2012).

Testing error: refers to the general error, and is the prediction error over an independent sample (Hastie *et al.*, 2014).

Training error: the average loss over the training sample (Hastie *et al.*, 2014).

Voluntary contributions within online forums: refer to voluntary knowledge sharing within online forums.

Webnography: or “virtual ethnography” (Morton, 2001). A spectrum of webnographic approaches exists from purely observational to full participatory (Kozinets, 2006).

Chapter 1 Introduction

1.1 Research context

The internet has allowed for massive amounts of knowledge to be made available online. Some providers of online knowledge are able to monetize their knowledge by charging for it. For example specialised journals and financial information services frequently charge for access to parts of their knowledge base. This study is concerned with online knowledge that is made freely available by individuals with no intention of direct financial reward (Wasco *et al.*, 2009). This thesis specifically studies online forums where individuals voluntarily seek and post information, typically on specialised subjects.

An online forum is defined as an interest-oriented online community based on the Web 2.0 technology that deals primarily with knowledge sharing by gathering members who share a common interest or theme (Spaulding, 2009). An online forum is firstly characterized by new technologies, i.e. Web 2.0, which has fostered a growing number of internet users (Mazurek, 2009). Creese (2007) suggests that the fundamental element of Web 2.0 is the collective action that can be reflected from the online knowledge sharing behaviours undertaken by the members of online communities. Collective action is associated with voluntary contribution behaviours (Ostrom, 2000). Members contribute knowledge around a topic or theme because they have common interests (Wasco *et al.*, 2009). For example, Dell online forum (www.support.Dell.com) is the Web 2.0 technology supported online platform where customers exchange knowledge about the usage of the product. Online forums generally involve some form of moderation and rules about who can share knowledge. However, the public often have access to all shared knowledge of the forum (Wasco *et al.*, 2009).

Online forums have been playing an increasingly important role in business (Füller *et al.*, 2008; Dholakia *et al.*, 2004; Demange, 2010), because they have emerged as a new interface with customers. For example, in generating new product ideas and resolving queries that might otherwise only have been resolved directly by the company's own (paid) employees (Demange, 2010). There has been extensive recent theorising in the marketing literature of “co-creation of value” between companies and their

customers (Vargo and Lusch, 2008a), and online forums represent a practical example of traditional producer-consumer boundaries being broken down with novel forms of value emerging.

This thesis distinguishes between information and knowledge that is shared in online forums. Knowledge is constructed through meaningful learning and occurs when the learner deliberately seeks to relate and incorporate new information to previously acquired knowledge (Braf, 2002). When an individual posts a message to an online forum, the post may represent a single piece of information, or previously synthesised knowledge of the poster. However, in posting, it contributes to greater collective knowledge. It is therefore understood that the collective knowledge available within an online forum emerges through individual contributions of information (Weinberger, 2007).

A significant challenge facing online forums is the continuing availability of knowledge from contributors, who by the definition of an online forum are not directly rewarded financially for their contributions (Harris and Rae, 2009). Digital knowledge illustrates the characteristics of a public good and is associated with a form of social dilemma (Wasco *et al.*, 2009; Riding and Wasco, 2010). Online forums provide intelligence of a collective, whose usage will neither exclude nor diminish the capability of access or usage by other users who follow (Wasco *et al.*, 2009). However, individual decision making models suggest that individuals' decisions are made to optimize a preference function subject to informational and material constraints (Gintis, 2007). By this approach, when every individual is rational and enjoys a public good for free, the public good will never be made (e.g. Ostrom, 2010).

For instance, an individualist member of an online forum knows he or she will not be excluded from the benefit of knowledge provided by that online forum, regardless of his or her own contribution of knowledge. However, the contribution of knowledge to online forums may cause potential risks such as time loss, anxiety or discomfort rose after disagreements. If every member of an online forum is purely individualist and selfish, there will be no valuable knowledge within the online forum (e.g. Wasco *et al.*, 2009). That is, there should be voluntary contributors who will pay for the cost of public goods in order to solve the social dilemma problem. Members of an online forum who benefit from the forum but who

wish to make little or no contribution to knowledge are referred to as free riders (Wasco *et al.*, 2009); Members who contribute knowledge to online forums are the voluntary contributors.

The term of “sustainability” has been used in ecology science and has a focus on environmental concerns, but it is now applied for others fields, such as for sustainable business (e.g. *Harvard Business Review*, 2008 <https://hbr.org/2008/09/harvard-business-ideacast-111>). Although there is not a common agreed definition for the term “sustainability” (Hoffman and Bazerman, 2007), it is often seen as a holistic approach involving aggregated factors that enable sustainable development (e.g. Sheth *et al.*, 2011). “Sustainability” is the goal and “sustainable development” is the dynamic process influenced by the numbers of identified factors to achieve the goal of “sustainability” (e.g. Diesendorf, 2000). The Cambridge dictionary defines the sustainability of something as “the ability to continue at a particular level for a period of time” (<http://dictionary.cambridge.org/fr/dictionnaire/anglais/sustainable>, 2016). The sustainability of online forums is understood in this thesis as aggregated factors that enable knowledge continuance within online forums over time. In order to be sustainable, forums need the constant availability of knowledge provided by users (Harris and Rae, 2009; Wasco *et al.*, 2009).

1.2 Research aims and objectives

This apparently circular process between consumption and contribution of knowledge in online forums has not been extensively researched. In particular, there is a lack of a complete understanding of the mechanisms that encourage individuals to voluntarily contribute knowledge to online forums. It becomes important to understand the reasons why some forums survive into a sustainable state, while others decline. The main research question of this thesis can therefore be summarised as, *how are online forums sustained?*

Existing studies tend to focus on isolated or limited numbers of antecedents of intention to contribute online forums. For example, studies examining antecedents of online voluntary contributions from a social perspective, refer to social recognition (Wasco and Faraj, 2005; Chiu *et al.*, 2006; He and Wei, 2009). Social recognition is an antecedent and outcome of interpersonal trust that impacts on the quality

of the dynamic interactions between individuals, who update their assessment that others will not act opportunistically to their detriment (McKnight and Chervany, 2002; Ridings *et al.*, 2002). Other studies have addressed motivations based on trust in online communities (Chen, 2007; Hsu *et al.*, 2007; Zimmer *et al.*, 2010; Zhang *et al.*, 2010). In other words, institutional trust can act as a facilitator of contextual structures and technology (McKnight and Cheveny, 2002), also contributing to online knowledge cooperation. Another research stream has highlighted the perceived critical mass of a forum, explaining that the usage of a forum is influenced by the perceived size of membership (Shen *et al.*, 2013). Studies address the success of new technology adoption on the basis of network size, arguing that individual followers take account on the perceived numbers of users of a system and decide whether to accept or reject that system (Roger, 1995). With regarding to the context of online knowledge contribution, perceived critical mass adds a social pressure to members to participate in online discussions (Shen *et al.*, 2013).

These studies have provided understanding of the constructs that have predictive effects on consistent online knowledge contribution behaviours. However, this only represents a minority of activities in investigating the motivational factors of online knowledge contribution (Chen, 2007; Zhang *et al.*, 2010). As a result of which, more empirical studies are needed. Although there is evidence that the perceived network size may facilitate online knowledge contribution (e.g. Shen *et al.*, 2013), the role of perceived critical mass has not been tested within an integrative model which considers others factors. Few studies have taken an integrative view in order to understand complex ongoing knowledge sharing behaviour (Hashim *et al.*, 2012). The first research question asks, *how do the key antecedents act together to influence online contribution behaviours?*

To answer this first research question, by testing an integrated predictive model, the study one develops a comprehensive understanding of the reasons that lead individuals to contribute knowledge to online forums. It firstly integrates members' perceptions of forum size with other previously discussed antecedents from a sociological perspective, namely trust in members and trust in online forums. It

identifies the mechanism and the patterns of online intentional behaviours, and tests the model with real data.

The study further looks at the interactions between antecedents and their importance to online knowledge contribution. Online forums involve fellow members and communities, therefore, trust studies in this area can also be divided into trust in members and trust in online forums. Trust in members has two levels: individual and interpersonal. Members may trust one particular member who can provide valuable ideas, but in most cases they communicate with the collective entities. In contrast, trust in online forums is institutional trust. Members may regard online forums as organizations or institutions, evaluating the mechanisms and conditions that online forums can provide. Although previous studies have distinguished the concepts of online, interpersonal (Wasco and Faraj, 2005) and institutional trust (Hsu *et al.*, 2007; Zimmer *et al.*, 2010), it is not clear how these different levels of trust together influence (or not) online knowledge continuance. The first investigative question in responding to research question one is, *how do the different levels of online trust impact on members' willingness for ongoing online knowledge sharing behaviours?*

Previous studies have sought to explain the success of online communities on the basis of network size, arguing that a tipping point may exist beyond which increasing economies of scale improve the chances of sustainability (Westland, 2010). A growing and successful forum may achieve higher levels of knowledge contribution once it has reached a critical size (Wasco *et al.*, 2009). This argument describes the consequence of online knowledge contributors' perception of network size. The causal relationships between perceived critical mass and online trust have not been examined previously, because the role that perceived critical mass plays in online knowledge contributions has only been tested with nested models (e.g. Shen *et al.*, 2013). Thus the second investigative question related to the research question one is, *how does perceived critical mass interact with the different levels of online trust?*

Identifying the mechanism or more general pattern that underlies the relationships among the antecedents of online intentional behaviours will enhance the predictive power of the integrative model. One

distinguishable characteristic of the causal model from the path analyses is that it is more stable in the long run and plausibly across different populations (Allison, 2014). The first contribution of this thesis leads to the examination of reasons for individuals' voluntarily contributing knowledge to online forums, exploring the causal relationships between motives for online contribution.

While previous studies have tended to take a static approach to modelling antecedents and consequences, a more dynamic model may be more appropriate. This leads to the second research question: *how does online trust evolve over time so that sustainable online forums can be attained?* Social influences in the study refer to online forum members' perceptions of their quality relationships with the forum and other members of the forum. For example, the level of trust they have in others forums members. Trust is multidimensional and different dimensionalities of trust have been addressed with previous studies from various fields (e.g. Larzelere and Huston, 1980; Ridings *et al.*, 2002; Lu *et al.*, 2009). Thus investigative question three follows from research question two, and is, *what are the dimensions of trust in the context of online forums?* The study also seeks to extend knowledge by examining the role of online trust in the continuance of supplying online knowledge, and to investigate how it is developed and undermined over time. This leads to the investigative question four: *how do the individual dimensions of trust contribute to overall trust development within online forums?*

The concept of perceived critical mass and the theory of critical mass are associated (Cho, 2011; Shen *et al.*, 2013). The theory of critical mass borrowed from percolation theory in Physics and describes how a mass phenomenon is evoked by a small group of members after a critical point (Oliver *et al.*, 1985). Marwell *et al.* (1988) argue that there is a possible self-reinforcing system in collective actions, and they highlight that this system is marked by group heterogeneity and an accelerating production function. Online knowledge contribution is a typical collective action (Wasco *et al.*, 2009). Collective action refers to the voluntary cooperation between groups of members in order to achieve their common goals (Meinzen-Dick *et al.*, 2004), for example, replying to members who seek knowledge. Answers to posts are often value-added because members can extend knowledge contributed by the previous repliers

(Wasco *et al.*, 2009). Group heterogeneity implies a small group of critical mass members who are either highly interested or have more resources to contribute, and will pay for the cost of the collective actions; the accelerate production function that allows the outputs of collective actions to exceed that by an individual, hence attracting more followers.

The emerging critical point and phase transitions are the key points implied by the theory of critical mass. Studies of the theory of critical mass are often conducted with network analysis in which the structural influences of the networks on the individuals' behaviours are highlighted (Westland, 2010; Centola, 2013). However, due to methodological constraints, it is practically impossible to measure the critical point in survey based studies. The concept of perceived critical mass can be used from sociology perspectives (Cho, 2011) to capture individuals' subjective perception of the critical phenomena after the critical point and their sequential behaviours are influenced by that perception (Shen *et al.*, 2013).

Thus, to fully explore the dynamic process within online forums, studying the concept of perceived critical mass alone cannot be satisfied. In fact, online forums can be seen as networks with nodes (members) having edges (relationships) between them. An edge links pair members when they communicate with each other such as contributing knowledge to an enquiry. Today, the possibility to access a big data set is increased and the methods to analyse and visualize online forums are developed in network science. However, knowledge from network analyses is rarely incorporated into human, economic and biological science (Barabasi, 2009). Westland (2010) argues that little research has sought to define the matrix for the theory of critical mass in order to understand what has happened before, near and after the critical point. Centola (2013) studies the emergence of critical mass members within the random networks, and his study inspires further research in this field.

However, the functionalities of networks vary with the structure of networks (Newman, 2005). Barabasi and Albert (1999) use the concept of "degree distribution" to explore the structure of a network. Degree distribution measures the structural pattern of relationships, and is particularly useful for understanding the functionalities and development of networks (Barabasi and Albert, 1999). Another approach defines

that the structure of online forums is measured through the size of memberships and the volume of message exchange because they reveal the intrinsic value of online forums (Wasco *et al.*, 2009; Ridings and Wasco, 2010). The second approach is helpful for discovering what has happened after the critical point. Studying the differences in the degree distribution (also the connectivity among members) is the major method to distinguish the type of networks (Newman, 2005), with their degree distributions mainly varying in Poisson, lognormal, stretched exponential and power-law (Clauset *et al.*, 2009).

Erdős and Rényi (1960) develop random network theory within which a member chooses to randomly connect to another with a given probability. The Poisson distribution has been used to describe the connectivity within random networks, informing that the vast majority of members have roughly the same connections and a few members have connections that deviate significantly from the average (Barabasi and Frangos, 2014). Randomised growth can also be described with the lognormal distribution. For example, the size growth or shrinkage of an organism depends on a random variable with its log10 converging to a normal distribution (Mitzenmacher, 2004). Another distribution in the network analysis can be the stretched exponential distribution which seeks to study the failure rate in a growth system; for example, email spam within a college (Newman, 2005). Although those distributions are highly peaked and skewed, they have finite variance around a mean value, i.e. the average connections (degree) associated to members.

However, a non-negative variable that follows the power law distribution does not fit this pattern (Clauset *et al.*, 2009). Members who are associated with a particular member within a network are called first-order of neighbours; neighbours of the first-order of neighbours are the second-order of neighbours (Barabasi and Frangos, 2014). The power law distribution that refers to variations between the numbers of the first and the second neighbours tend to be infinite. There is not a typical scale to describe such variance, thus it is named scale-free (Barabasi and Frangos, 2014). The concept of scale-free suggests a power law tail or power law with the exponential cut off characterising the distribution of membership connectivity marked by growth and preferential attachment (Newman, 2005). For example, a web grows

in time by attracting other links connecting to it, and the new jointed links prefer to attach to a web depending on the previous level of links that the web has (Barabasi and Albert, 1999).

In scale-free networks, the majority of connections of a network are held in a small group of hubs which have higher connections, while the vast numbers of others have few or zero connections (Barabais and Albert, 1999). Hubs therefore have a similar meaning with the critical mass members, who should have more connections because they contribute more knowledge than the free riders. This characteristic also makes a scale-free network vulnerable to attending attacks on the hubs (Barabasi and Frangos, 2014), which leads to an understanding that the critical mass members play an essential role in sustaining online forums. Another important functionality that distinguishes a scale-free network is that it is self-sustaining (Cohen *et al.*, 2002; Barabasi and Frangos, 2014; Newman, 2005). Drawn from percolation theory, studies of the cluster structure show that a spinning cluster always exists within a scale-free network, and it always percolates (Barabasi and Frangos, 2014). A spinning cluster that allows the network to be connected by hubs is associated with an infinite system, and occurs almost immediately after a critical point (Cohen *et al.*, 2002; Barabasi, 2013; Newman, 2005). In other words, there are always newly joined members attaching to hubs who become more popular over time (Newman, 2005). Existing studies have sought to understand under what conditions the spinning cluster and hubs emerge (e.g. Cohen *et al.*, 2002). Similarly, the study of the theory of critical mass applied within networks has a focus on the initial conditions for the critical mass members being in existence (Centola, 2013). On the basis of the above discussions, the structure of online forums should be firstly examined in order to explore the role of the online critical mass contributors.

To fully explore the dynamic aspects of critical mass, structural influences in this study refer to formal structure within a forum, and quantified rather than qualitative relationships between members. This leads to the following research question three: *how can the theory of critical mass be applied to understand the structural influence of online forums in relation to knowledge contribution continuance?* This study seeks to extend knowledge of critical mass theory by examining its formal structure in the context of online

forums, and how this affects ongoing online knowledge sharing behaviours. To answer research question three, the investigative questions five and six are asked accordingly: how is the critical point beyond which the mass phenomenon of knowledge sharing within online forums achieved? What happens before and after the critical point in terms of online knowledge contributions?

1.3 The present research

Previous research has found that online trust and perceived critical mass are antecedents of online knowledge continuance. Little is known about how dynamic factors combine together to have an effect on ongoing online discussions, and how those dynamic antecedents act in relation to each other. Thus, study one firstly seeks to propose a conceptual model that combines the antecedents of continuous online knowledge exchange behaviours. The multi-levels concept of online trust is identified as the key social factor. Perceived critical mass as the structural influences is the other key antecedent. The dynamic aspects of online trust and critical mass are further examined in study two and study three accordingly.

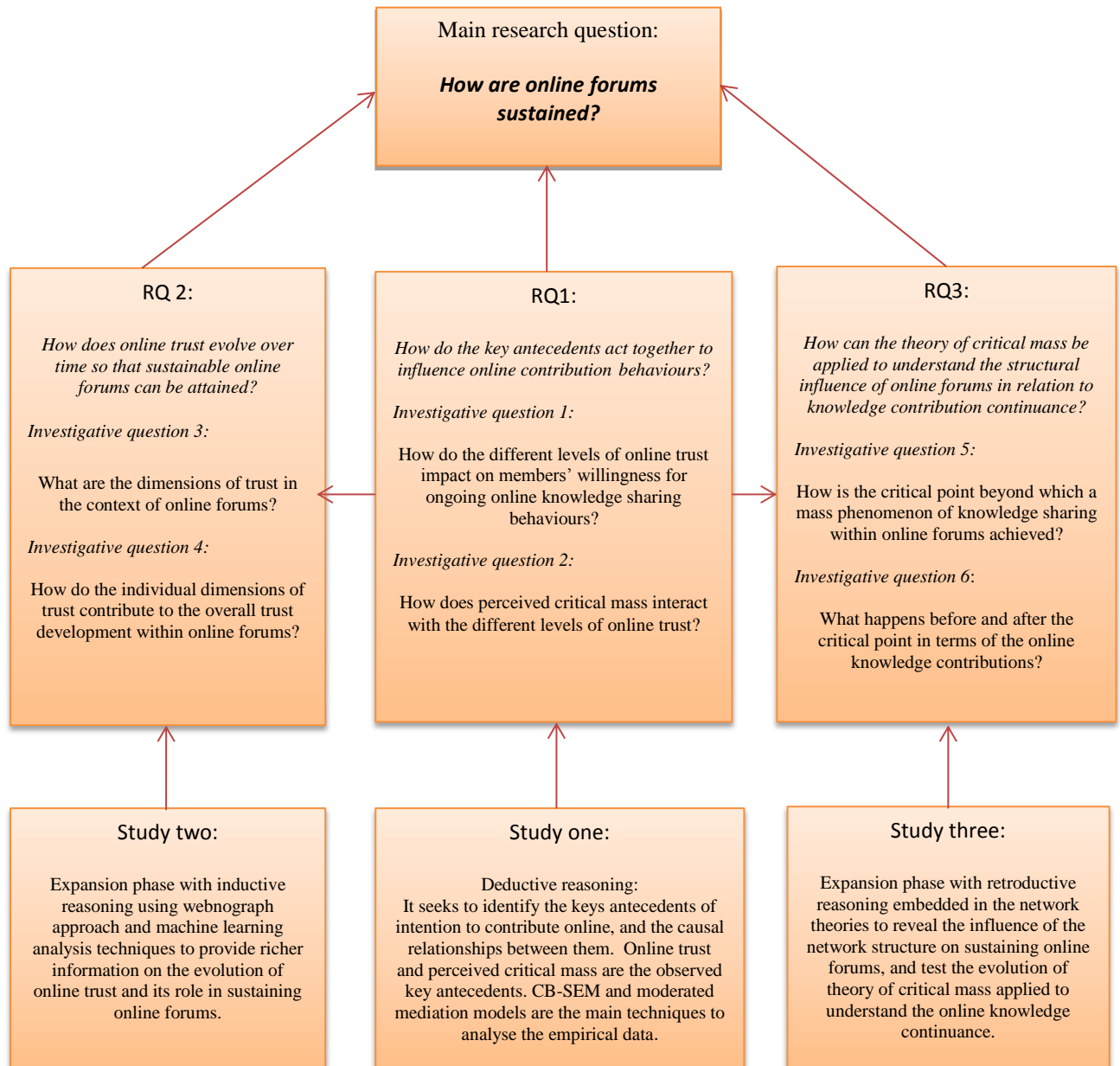
The proposed conceptual model is embedded in the decomposed theory of planned behaviours (DTPB) (Taylor and Todd, 1995) which is an extension of the theory of planned behaviour (TPB) (Ajzen, 1991). TPB suggests that intention can lead to actual behaviour (Ajzen, 1991). Future-orientated intention can reflect an individual's strong willingness and motivation to perform behaviours (Bratman, 1984). It is therefore argued that future-orientated intention is associated with ongoing intention to share knowledge online, without which online discussion cannot be dynamic. Although TPB claims to be a generalized model that has been widely applied, specific antecedents vary according to the context (Taylor and Todd, 1995). To overcome this limitation, DTPB decomposes the antecedent components in TPB in order to provide greater understanding of the intentional behaviour in a particular context (Taylor and Todd, 1995). Following the TPB logic, DTPB provides a more precise understanding of the determinants than TPB (Chennamaneni *et al.*, 2012). For instance, DTPB can be developed for understanding knowledge sharing within organisations by decomposing the attitudinal beliefs into perceived organisational incentives, perceived reputation enhancement and perceived enjoyment in helping others (Chennamaneni

et al., 2012). DTPB decomposes the determinants of TPB in order to satisfy the variations of different contexts under study, in this case, online discussion forums that should be distinguished from traditional communities within which members often have physical contacts. The emphasis of DTPB is not to test the immediate determinants of intention, but in exploring the contextual factors of these determinants (Taylor and Todd, 1995). DTPB is introduced in study one to allow behavioural, normative and control beliefs to be decomposed into multiple dimensions to fit the specific context (Rogers, 1995) of online forums.

Study one also seeks to examine the causal relationships between these key antecedents that are dynamic in nature in order to enhance the predictive power of the proposed model. The combined effects of identified antecedents are tested with moderation and mediation analyses following the logic proposed by Baron and Kenny (1986) and Preacher *et al.* (2007).

Study one serves to provide a general picture by including the key antecedents of online cooperative behaviours. It is designed to answer the research question one about the reasoning for online knowledge contributions. Although the analyses of causal relationships among the identified dynamic antecedents are able to offer a richer understanding about how these antecedents impact together on online knowledge continuance, the dynamic nature of the antecedents is not able to be fully explored. To address this limitation of study one, two studies are conducted following the proposed conceptual model. Study two seeks to answer the research question two, regarding how online trust is developed over time. The research question three, with respect to how perceived critical mass is associated with the structural dynamics that impact on the ongoing online contribution behaviours, is addressed in study three. That is, this thesis is composed of three separate yet connected studies (see figure 1).

Figure 1: Rationale of this thesis



The rationale of this thesis is as follows. Firstly, it discusses the nature of online voluntary contribution and antecedent factors that have been associated with individuals' propensity to contribute knowledge to online forums. The study specifically focuses on trust in members, trust in online forums and perceived critical mass. These antecedents are integrated within a framework based on decomposed theory of

planned behaviour. From the literature, it identifies the gaps in knowledge and specifies hypotheses for testing with real data from an online survey, using covariance-based structural equation modelling and moderation/mediation models. Secondly, it analyses the dynamic nature of online trust in a brand with content left by reviewers from three different online forums. Machine learning analysis techniques are used to distinguish the dimensions of online overall trust, and to examine the evolution of each dimension contributing to online trust building and undermining. Thirdly, it evaluates the role of the critical mass theory applied in online knowledge contribution continuance, and examines the mechanisms that govern the development of the online forum being seen as a network. Network analyses are embedded in the data collected from a large size online forum.

In study one, 910 responses are collected through an online survey to examine the conceptual model that identifies online interpersonal trust, online institutional trust and perceived critical mass as key determinants of the antecedents for continuous intention to contribute online, and the causal relationships between antecedents. Covariance-based structural equation modelling (CB-SEM) (Wright *et al.*, 2012) is the main analysis technique employed in study one. Following the logic developed by Baron and Kenny (1986), eleven moderation/mediation models are created to understand how the identified antecedents are mutually influenced together on ongoing intentional online contribution behaviours.

Results generated from study one confirm most of the hypotheses developed in the conceptual framework. Ability and benevolence-based trust in online forums positively impacts on the attitudinal and behavioural control intention. Perceived linkage to the critical mass members and many numbers of contributors within online forums are essential in normative intention. Trust in members who are not behaving opportunistically favours the development of trust in online forums, but does not impact on the perceived mass usage of online forums. Results from the mediation analyses show that the positive effects of trust in members to the attitudinal intention are completely mediated and moderated by trust in online forums that are able to and would like to allow many communicators to contribute within online forums. Moreover, the effects of trust in members impacting on normative intentions are mediated by perceived

critical mass. However, such mediation effects can disappear with the introduction of trust in online forums. Thus the role of trust in members cannot be excluded, but the role of institutional trust is highlighted. Taken together, results show that the three identified antecedents affect the online intentional contribution behaviours in different magnitude levels, in which trust in online forums has more weighted power in the prediction of the online intentional contribution behaviours.

Study two then seeks to understand the evolution of institutional trust and takes a webnographic approach. Themes are developed based on an analysis of 1131 comments left by reviewers about Skype internet telephone communication service during the period 2005 to 2010. Support vector machine (SVM) is used to examine whether the developed themes can reflect distinct components of online trust. Neural network analysis (NNA) will also be incorporated to analyse the evolution of online institutional trust over time.

Results from study two shows that online institutional trust can be better understood through two dimensions: ability and benevolence. Results also show that “ability” plays an important role in the earlier stage of initial trust-building, but “benevolence” that is associated with emotions affects online institutional trust in the later stage. Results are in agreement with the measurement developed for trust in online forums in study one, that competence and benevolence are two significant dimensions of online institutional trust. However, results from study two provide richer understandings by informing the focus of online trust management in different stages, so that its role in sustaining online forums is possible.

Study three seeks to understand the role of critical mass members in sustaining online forums, and uses network structural analysis techniques proposed by Clauset *et al.* (2009), inferring a different methodology from SEM and webnography. An empirical study of a successful online forum “Stack Overflow” is undertaken to explore the structure of online forums. Analysis is performed on time-series data from 2008 to 2012, which produces a data set of 147190 nodes and 149289 edges.

The results of study three show that successful online forums are scale-free networks that in nature are self-sustaining (Barabasi and Albert, 1999). The network attack simulations show that it is the small group of critical mass members (around 5% of memberships) who sustain online knowledge contributions. The critical point using the technique proposed by Cohen *et al.* (2002) is calculated at 0.0078 closing to zero. The critical point over which online forums will be self-sustaining is quickly achieved, in agreement with the argument by Barabasi (2013) that a scale-free network always percolates. This explains that the concept of critical mass in the case of successful online forums can be understood through a sociological perspective, and by examining the phenomena after the critical point (Cho, 2011). However, far beyond the critical point, the connections to the initial emerged critical mass members tend to withdraw (knowledge exchanges between the critical mass members and other members are less intensive), while the size of the online forum keeps on growing in a steady state. This may be explained with the findings from study one that member participation in online activity is more likely to be influenced by their perceived size of network. This may also be explained through the production function proposed by the theory of critical mass, i.e. the total knowledge provided by the online forum exceeds the knowledge contributed by an individual. As a consequence of this, the initial contributor may not be able to contribute knowledge that can respond to the inquiries of every member, and the newly jointed members do not necessarily only attach to the initial contributors.

Despite the above discussions, contributors are encouraged to contribute knowledge online during the evolution of the online forum under study; in return, they can win more connections because newly joined members tend to proportionally attach to contributors embedded in their previous contribution. This is in agreement with the previous findings (e.g. Newman, 2005) that the structure of scale-free networks allows key contributors to become more popular over time, suggesting the structural influence on the online contribution behaviours. Results show that one initial contributor who has the highest degree is observed presenting far above the critical point. This can be explained with the argument by Barabasi and Frangos (2014) that the probability of obtaining more connections by the initial members is higher than that by the later emerged members in the case of the scale-free networks.

Contribution of knowledge within an online forum implies usage of that forum; it also can be inferred that ties are created among members (Haythornthwaite, 2002). Frequent reciprocal contacts over time between individuals are characteristic of ‘strong’ ties (e.g. Haythornthwaite, 2002) and they often involve trust between pairs (Krackhardt *et al.*, 1992). Trust in members at the individual level can reveal the strong relationships between them. In contrast, “weak” ties suggest that the probability of two untied parties becoming tied will increase when these two parties have common acquaintance(s) (Granovetter, 1973). “Weak” ties are used to describe the relationships within a broad communication system (Haythornthwaite, 2002). For example, members may have a sense of belonging to an online forum and trust their peers, but they are only tied by that online forum (Haythornthwaite, 2002). To use another example, members are tied because they have the common acquaintances (Rapoport, 1957). Thus, the concepts such as trust in members at the interpersonal level, trust in online forums and perceived critical mass may be associated with weak ties. Previous studies of computer-mediated communication (CMC) have been concerned with the relationships maintained via online social networks (Haythornthwaite, 2002). The “richness approach” is emphasized in the design of CMC and technology. The “social presence” approach ignores the strong ties but highlights the importance of weak ties among communicators (Haythornthwaite, 2002).

Both study one and study three find out how those strong ties can lead to weak ties in the context of online forums. Findings from study one show that trust in members at the individual level (strong ties) can promote a more general trust in online forums (weak ties). Findings from study three shows that scale-free networks are connected, suggesting that every member can be reached through the critical mass members. In other words, it is possible that every member knows each other (weak ties) (Haythornthwaite, 2002). Meanwhile, only a small group of members contribute knowledge within online forums (Wasco *et al.*, 2009), and it is the critical mass members who sustain the online knowledge contribution. Therefore, the communication pattern within online forums is marked by “few to many”, and the communications between this small group of members and others members are intensive (strong

ties). This is in agreement with Granovetter's (1973) argument that "weak" ties depend upon on how "strong" the ties that the untied parties create with their common acquaintances.

The communication pattern that suggests "few are strong ties but many are weak ties" can also explain the results from study one: why weak ties influence strong ties, because members trust in an online forum that is concerned about the ability of its members to get along (weak ties). They therefore tend to intensively communicate with a small group of particular critical mass members who are able to provide valuable knowledge within online forums (strong ties).

1.4 Principle contributions of thesis

Research on the sustainability of online forums is rare (Ridings and Wasco, 2010). This thesis, embedded in empirical studies, seeks to contribute knowledge in the field by taking different worldviews with the aim of providing a more complementary understanding of the topic. The different aspects of the world cannot be described through only a single worldview (Hes and Du Plessis, 2014). However, research that crosses disciplines in this field is a minority activity, and research incorporating a multiple worldview is not popular.

While previous studies provide relevant insights into the factors associated with contribution to online forums, they predominantly take an un-prophetic approach and do not capture the causal interactions among variables. Few studies have taken an integrative view in order to understand complex ongoing knowledge sharing behavior (Hashim et al., 2012). This thesis identifies the influence of both social and structural dynamics on the intention to share knowledge online, and examines the complex relationships between dynamic antecedents that are poorly understood (Ridings and Wasco, 2010). Understanding how these identified dynamic antecedents affect the online contribution behaviours is important, because this can enhance the predictive power of the proposed model. However, this consideration seems to be ignored in the prior studies.

Existing studies have little investigated the dynamic aspects of these antecedents. For example, online trust is often cross-sectional, and therefore the results generated from study two add to the literature on the dimensions of trust, and make a contribution by identifying how these dimensions change in salience over time (Palmer and Huo, 2013). Another example, critical mass is often studied through the measurement of the mass phenomena after the critical point (Cho, 2011); study three investigates what has happened before, near and after the critical point, as proposed by the theory of critical mass applied in the context of online forums.

For managerial considerations, it is suggested that firms who use online forums as internet media to communicate with consumers should beware of the process of managerial strategy development. The value of knowledge perceived by members plays an important role in the initial stages. In this stage, firms should be very concerned of how to get members who are knowledgeable to contribute online. Incentive means to attract experts can be used, because it is those initial contributors who will evoke the expansion of online discussions forums. This proposition is in agreement with previous findings that high knowledge quality will attract new followers (Chen, 2007) and promote word-of-mouth in the beginning (Libai *et al.*, 2013). At a later stage, subjective feelings on the good will demonstrated by firms are more important, because the results from study two show that benevolence has an important influential impact on the undermining of trust. With the numbers of users increased, the objective is to allow the hosted online forum to be scale-free. A scale-free network is self-sustaining, and it allows the contributors of knowledge to become more popular over time (Newman, 2005), i.e. the critical mass members are delicately maintained because of its structural influences. Finally, the technology support should be provided during the whole process of development. For example, thousands or millions of members online simultaneously; limited moderation on the exchanges within members; voting system that encourage members to publish quality content and so on. This understanding requires niche strategies that can respond to the altered focus in the different stages of online trust building. By doing so, firms can effectively manage online forums and reduce associated cost.

1.5 Limitations

This research is limited in different ways. Study one is cross-sectional, in which the dynamic aspects of online trust and critical mass can only be measured through an online survey. Although study two investigates the evolution of online trust, it is not a longitudinal study, but rather a time series approach, as it is almost impossible to record the comments left by a particular reviewer over 6 years. Study three identifies the position of critical mass members through only one centrality measure. Different centrality measurements may lead to different results. Other indexes of the importance of members, such as the diffusion-based algorithm are not tested in the context of online forums. For future research, it is recommended that there will be more longitudinal studies in the sustainability of online communities that takes different worldviews.

1.6 Thesis outline

This thesis is composed of seven chapters, two appendices and bibliography. Chapter 1 introduces the research background related to the topic under study, and seeks to explain how this thesis could add knowledge to existing studies.

Chapter 2 reviews the theoretical background of the antecedents to the determinants of intentional online contribution behaviours. This involves three main perspectives: social research, network analyses and an interdisciplinary approach. On the basis of these theoretical views, a conceptual framework is firstly developed, which integrates the identified antecedents so that an overall outlook on the factors influencing the online knowledge continuance is provided. The research hypotheses are proposed and an integrative model is presented. Followed the conceptual framework, the dynamic aspects of the identified antecedents are further discussed in chapter 2.

Chapter 3 explains a mixed-method research design. A deductive approach is served to test the hypotheses developed in chapter 2, represented by study one. The results from study one are further expanded through an inductive study two and a retroductive study three. The analysis techniques associated to each study are also discussed in chapter 3.

Chapter 4 shows the results from study one. It describes the key features of the preliminary and the main study analysis. The preliminary analysis includes the descriptive data, non-response bias assessment and the sample file; the main study essentially presents the results from covariance-based structural modelling and the mediation / moderation analysis.

Chapter 5 presents the results from study two. Firstly, it shows the results of the coding, within which four themes represent the trust emerging dimensions. Secondly, the results from the analyses that tend to examine the distinctness of these four emerged dimensionalities are presented. Thirdly, it describes how online trust evolves over time.

Chapter 6 displays the results from study three. It firstly discusses the results that suggest the online forum under study is a scale-free network; hence the network topology with regarding to this type of network is able to be applied. On this basis, the critical point highlighted in the theory of critical mass is calculated, which informs that the evolution of the online forum can be analysed using four successive stages (i.e. far before, near, above and far above the critical point). Followed this, the results about the mechanisms that govern the evolution of the online forum in different stages are discussed. The influence of the critical mass members during the development of the online forum are also explained in chapter 6.

Chapter 7 further concludes the theoretical contributions and the managerial considerations of this thesis, which are embedded in the findings from the three studies. It answers to the research questions and associated investigative questions proposed in chapter 1.

Finally, the bibliography and appendices are presented. The appendices comprise the online survey questionnaire used in study one, and the coding from study two.

Chapter 2 Literature review

This chapter starts by defining online forums. The online social dilemma is recognised as one of the biggest challenges to sustaining an online forum. Knowledge contribution which is an online collective action is identified as a solution to the online social dilemma. To understand how online knowledge contribution is instrumental in sustaining online forums, it is necessary to consider the motivational dynamics of online contribution behaviours.

Given the complex nature of the topic undertaken, the study takes a cross-disciplinary approach by using three studies to answer the main research question. Study one seeks to develop hypotheses within an integrated predictive model that is embedded in the decomposed theory of planned behaviour (DTPB) (Tylor and Todds, 1995). The dynamic constructs are identified, and the causal relationships between them are investigated in order to understand how those antecedents interact with each other and have an effect on ongoing online discussions. The dynamic aspects of the identified antecedents are discussed separately in study two and study three. Section 2.2 discusses the importance of the topic undertaken by the thesis, raises the research questions relating to the theoretical gaps, and presents the rationale of using multiple studies in the thesis.

The theoretical backgrounds of this thesis are deliberated in section 2.3. Trust in members, trust in online forums and perceived critical mass are identified as the key dynamic antecedents of online voluntary contributions. The conceptualizations of trust and critical mass are discussed in sections 2.3.1. Section 2.3.2 develops hypotheses within the integrative model. The dynamic aspects of trust within the context of web-based communications are presented in section 2.3.3. In order to fully explore the concept of critical mass within online forums, related theories from network science are considered in section 2.3.4.

2.1 Introductory elements

There is now considerable evidence that online forums play an increasingly important role in business and society (Füller et al., 2008; Dholakia et al., 2009; Demange, 2010). The use of online forums is consistent with firms' efforts to encourage customer "co-creation" of value (Vargo and Lusch, 2008). By developing an online forum, a firm can facilitate its customers to create value by sharing knowledge with the company and fellow customers, and thereby also reducing its own costs. What is the definition of online forums? Why is it important for business and society? The following sections 2.1.1 and 2.1.2 seek to answer these two questions, and to explain the importance of undertaking the research which forms the basis of this thesis

2.1.1 Definition of online forums

Online forums are interest-oriented online communities (Spaulding, 2009). The definition of online communities is associated with the term Web 2.0 (Mazurek, 2009). This section discusses the concepts of Web 2.0, online communities and online forums.

Marketers have long recognised the power of promulgating compelling messages with the objective of changing attitudes and behaviour. Over past decades, studies of advertising and promotion effectiveness are copious. However, firms cannot depend exclusively on traditional tools in order to stimulate consumption (Palmer and Koenig-Lewis, 2009) for two main reasons. Firstly, consumers are becoming increasingly savvy about products or services. Secondly, if, manipulation on the part of firms is too overt this can lead to negative attitudes and resistance to repeat purchase (Krishnamurthy and Kucuk, 2009). Most recently, this challenge to direct marketing means that firms now have to find alternative channels for communicating with consumers.

There is currently a lot of discussion concerning Web 2.0 techniques within marketing researchers. For example, the phase Research 2.0 appeared in 2006 to cover various methods based on online collaborative techniques (Comley, 2008). Financial services have used Web 2.0 and Customer 2.0 with the aim of staying customer-focused and maintaining trust (Stone, 2009). According to the McKinsey Quarterly's

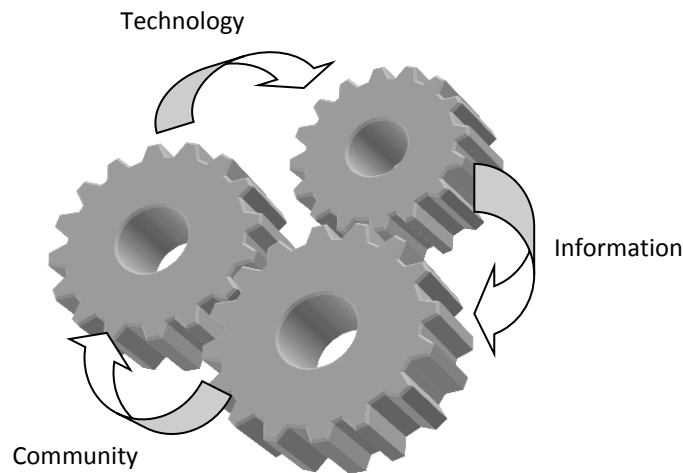
survey on Web 2.0 in 2013, 83 percent of enterprises in their sample claimed that their business obtains measurable gains by implementing Web 2.0 techniques (Mckinsey Quarterly's survey, 2013).

The term Web 2.0 has been understood from different perspectives. Web 2.0 is an combination of existing protocols and computer languages, such as peer-to-peer technologies, semantic web, RSS (Really Simple Syndication) feeds as well as AJAX (Asynchronous JavaScript and XML- eXtensible Makeup Language)(Kim *et al.*, 2009). Web 2.0 reflects a technology push that enables the proliferation of blogs, wikis, RSS feeds, peer-to-peer networks and so on, to allow information and knowledge sharing computing systems (Kim *et al.*, 2009).

In contrast to this purely technical approach to Web 2.0, Creese (2007) suggests that the fundamental element of Web 2.0 is the collective action that can be reflected from the online information sharing behaviours undertaken by the members of online communities. O'Reilly (2005) argues that a good indicator of success of a new technology can be the increasing number of usage. Similarly, Weinberger (2007) describes Web 2.0 as the people oriented virtual environment, because internet users play a more important role in the web world than they did one decade ago (Wu and Yang, 2009).

In summary, Web 2.0 is characterized by new technologies that have fostered a growing number of internet users who share information within online communities (Mazurek, 2009; Wu and Yang, 2009). The term Web 2.0 is applied to differentiate its new features notably "shared value" from previous Internet communication (Högg *et al.*, 2006). "Shared value" means that the information flows within an online community are characterized as "many-to-many" communication pattern, which is contrary to the "one-to-many" one decade ago (Wang *et al.*, 2008; Karakas, 2009; Kim *et al.*, 2008). Web 2.0 technology allows massive online communication and is an improved technical architecture that empowers computer usage (Weinberger, 2007). Figure 2 summarizes the different perspectives of Web 2.0 concepts.

Figure 2: Three different perspectives of Web 2.0 concepts



Source: Grzegorz MAZUREK (2009, p.70). Three different perspectives of Web 2.0 concept, Web 2.0 implication on marketing, management of organizations.

The word community has two Latin derivations (*Oxford English dictionary*, 2000): community, used in the context of ‘common fellowship, society’; co (m) munet, applied in the context of ‘fellowship, community of relations or feelings.’ Since the medieval Latin period, this word has extensive meanings such as “universitas”, which reflects “a body of fellows or fellow-townsmen.” To date, the term “community” has a wider usage, along with the concept of the general public interest: “the community: the people of a country (or district) as a whole; the general body to which all alike belong, the public” .

Kozinets (1999, p. 253) defines a community as “a group of people who share social interactions, social ties, and a common place”. According to Bell and Newby (2004), the key contexts of community refer to social interaction, geographic area, and common bonding. According to De Moor and Weigand (2007), the success of a community relies largely on the strong and lasting social interactions that occur in some form of common space.

Plant (2004) describes an online community as “a collective group of entities, individuals or organizations that come together either temporarily or permanently through an electronic medium to interact in a common problem or interest space” (p.54). Different to traditional communities, online communities are larger in space and time, since the physical location is irrelevant. In other words, participants may come from different countries, regions or cultures, with the result that members of online communities are more heterogeneous with respect to social characteristics such as lifestyle, ethnicity and socio-economic status (Ridings *et al.*, 2002).

Online communities comprise many forms of communication, such as blogs (Bishop, 2009), electronic social networks (Clemons *et al.*, 2007), and online forums (Brown *et al.*, 2007). Previous research has classified online communities into four main categories (Kannan *et al.* 2000; Spaulding, 2009):

1. Transaction-oriented communities, such as eBay.com, which emphasize processes of online transactions of products where buyers and sellers must interact to complete transactional information;
2. Relationship-oriented communities which encourage members to provide accurate and authentic personal information to build relationships, such as the family or friend relationships (e.g. facebook.com) and business relationships (e.g. LinkedIn.com);
3. Fantasy-oriented communities which gather members who play out their fantasies in the virtual world, such as the Second life which provides entertainment, playfulness and joyfulness to members (Clemons *et al.*, 2007);
4. Interest-oriented communities which deal primarily with knowledge and information sharing by gathering members who share a common interest or theme, such as Dell online forums (www.support.Dell.com) where members exchange information about the usage of the product.

This research focuses on interest-oriented online forums. Online forums are defined here as web-based platforms which allow people to share knowledge. They generally involve some form of moderation and

rules about who can share knowledge, initiate new “threads” and have access to all or parts of the forum. However, publics often have access to all shared knowledge of the forum. An online forum may be linked to a commercial organisation whose products are the focus of the forum (e.g. a software company that sets up a self-help group of users). However, the firm’s interest will not be primarily transactional, in other words, there is no direct exchange of money in return for knowledge contributed or consulted. Technically, online forums sit between older format bulletin boards and more modern “groups” within web-based media platforms. The focus of this paper is the forum, but many of the principles discussed here could also apply to the older and newer technologies.

2.1.2 Social dilemmas in online forums

One challenge faced by a sustainable electronic forum is the availability of knowledge/ information (Chiu *et al.*, 2006; Harris and Rae, 2009; He and Wei, 2009; Levy, 2009; Payne *et al.*, 2009). Online information satisfies the conditions of the definition of a public good (Wasco *et al.*, 2009; Riding and Wasco, 2010). The four dimensions that characterize public goods are non-rivalry, non-excludability, the accelerating production function and jointness of supply (Wasco *et al.*, 2009).

Non-rivalry means that a public good such as a public park and public radio are not used up after consumption (Samuelson, 1954). Similarly, online shared information keeps the same content after being viewed, and viewing by one person does not diminish the capability of access or usage by others following after (Wasco *et al.*, 2009).

Non-excludability refers to all individuals in a collective having a right to access a public good, regardless of their contributions (Snidal, 1979). In the context of online forums, all members can benefit from contributed messages. However, it is noted that non-excludability should be distinguished from the inability to control exclusion (Snidal, 1979). The latter could be happened if moderators of an online forum remove the contribution of a specific member, or forbid a particular person accessing the contributions of others members (Wasco *et al.*, 2009).

The production function describes the unbalanced relationship between input and output. The input often does not have an influence on the output of a public good. For instance, the cost of constructing a public park is fixed once it is built regardless of the number of visitors. It is comparable with the cost of a published message in an online community which remains the same no matter how many members review that message (Wasco *et al.*, 2009).

The cost of joint production of public goods is lower than would be the case if the good was produced separately (Buchanan, 1966). A classic example is theatre performance, which through television or digital broadcast, individuals can see at home. In parallel, it takes more resource for an online forum to develop software to codify the first visitor than members who join after (Wasco *et al.*, 2009).

The classic linear public good is illustrated through the utility function: $U_i = U_i[(E - X_i) + A * P(\sum X_i)]$, where “E is an individual endowment of assets; X_i is the amount of this endowment contributed to provide the good, A is the allocation formula, and P is the production function. A is specified as $1/N$ and $0 < 1/N < P < 1$, where N is the number of individuals (Ostrom, 2000, p.139)”. Regarding to the digital public goods, it is indivisible, i.e. the information online remains the same however many times it has been read. Thus, the utility function of the digital public goods is expressed as $U_i = U_i[(E - X_i) + P(\sum X_i)]$, with $0 < 1/N < P < 1$.

One obvious problem associated with public goods is the social dilemma that will occur when the public good has both non-rivalry and non-excludability (Ostrom, 2000; Ostrom, 2010). According to Gintis (2007), the individual decision making model (named as a rational actor model in economics) indicates that decisions are made to optimise a preference function subject to information and material constraints.

In other words, when every individual is rational and enjoys a public good for free, the public good is never produced. Similarly, members of an online forum who benefit from this network but who wish to make little or no contribution are called free riders. It has been observed that potential losses are associated with contribution to online forums. A contributor sacrifices their time participating in online

discussions. Members may submit their private information, and anxieties could arise about the ability of an online forum to treat such information confidentially. Members' opinions are influenced by others and social risk may be present in how peers evaluate a member's contribution (de Valck *et al.*, 2009).

Overcoming the public good/social dilemma problem requires voluntary cooperation (Ostrom, 2000; Wasco *et al.*, 2009). Despite incurring perceived costs which are not matched by any specific benefit, there is evidence that individuals may continue contributing to online communities (Cummings *et al.*, 2002). In this study, this behaviour is described as "online voluntary contribution", defined as the giving of knowledge by forum members who incur opportunity costs in contributing knowledge without specific expectations of material reward.

2. 2 Rationale of using multimethod approach

The sustainability of online forums is tightly associated with knowledge continuance (e.g. Wasco *et al.*, 2009). However, the existence of social dilemma can make sustaining an online forum very challenging. Previous studies have found out that the ongoing online voluntary knowledge contribution being the key for a successful online forum (Ridings and Wasco, 2010). Studies on the antecedents of mass voluntary contribution in online contexts are rare, although their presence may provide intrinsic reasons for the existence of online networks (Wasco *et al.*, 2009). Dynamic online discussions consist of evolving online voluntary contributions (Wasco and Faraj, 2005; Wasco *et al.*, 2009). This requires identifying constructs that are dynamic in nature.

In order to explore the dynamic aspects of identified antecedents, this thesis takes a cross-disciplinary approach with three studies. Study one seeks to investigate whether these dynamic antecedents play the role in online knowledge contribution behaviours. If so, how they act together in the context of online forums. The digital world can provide a rich time series data for scholars and firms. In this context, study two and three can address the limitation of the cross-sectional study and examine the evolution of antecedents identified with study one.

Previous studies have examined the dynamic antecedents of online voluntary contributions from a social perspective, including social recognition such as the reputation obtained by sharing knowledge with others (Wasco and Faraj, 2005; Chiu *et al.*, 2006; He and Wei, 2009), trust in online organizations (Chen, 2007; Hsu *et al.*, 2007; Zimmer *et al.*, 2010; Zhang *et al.*, 2010), and the perceived numbers of online contributors (Shen *et al.*, 2013).

Embedded in these studies, it is understood that perceived critical mass, interpersonal and institutional trust are essential influences on the online knowledge cooperation. Perceived critical mass suggests that individual followers are influenced by their perceived numbers of usages of that system (Roger, 1995). McKnight and Chervany (2002) define “reputation” as both the antecedents and consequences of interpersonal trust. Interpersonal trust refers to trust in others who will not perform opportunistically (Ridings *et al.*, 2002). Interpersonal trust therefore describes the quality of interactions within individuals. Institutional trust involves two sub-constructs (Mcknight and Cheveny, 2002): situation normality and structure insurance. Situational normality means that people are more likely to trust when facing explicable and normal situations (e.g. Gefen *et al.*, 2003). Structure insurance refers to the expected successes being facilitated by the contextual factors such as guarantees, regulations, technology and so on (e.g. Pavlou and Gefen, 2004).

However, few studies have taken an integrative view in order to understand complex ongoing knowledge sharing behaviour (Hashim, 2012). *How do the key antecedents act together to influence online contribution behaviours?* Although previous studies have distinguished the concept of online trust from interpersonal trust (Wasco and Faraj, 2005) and institutional trust (Hsu *et al.*, 2007; Zimmer *et al.*, 2010), it is not clear how these different levels of online trust influence (or not) online knowledge continuance. *How do the different levels of online trust impact on members’ willingness for ongoing online knowledge sharing behaviours?* Moreover, online knowledge sharing is characteristic of a general communication pattern (Wasco *et al.*, 2009), and the perceived mass numbers of members has a social influence on online knowledge sharing (Shen *et al.*, 2013). However, *how does perceived critical mass act in relationships*

with the different levels of online trust? Study one has sought to propose an integrated predictive model, and to extend knowledge by examining the multiple levels of online trust within online forums. Furthermore, this study will analyse the relationships between identified antecedents that act on each other.

The development of trust involves a gradual process to build (Li *et al.*, 2008), and to undermine (Charki and Josserand 2008). Although trust has been studied in various disciplines (Ferrin and Dirks, 2006) and is studied in different marketing contexts (Pavlou and Gefen, 2004; Luo, 2006), the empirical study of trust in computer mediated peer-to-peer environments remains limited (Vance *et al.*, 2008). Furthermore, most studies of trust have taken a cross-sectional approach (Kassim and Abdulla, 2006; Massey and Dawes, 2007; Gupta *et al.*, 2009), with relatively few time-series studies (Palmer and Huo, 2013). However, *how does online trust evolve over time so that sustainable online forums can be attained? What are the dimensions of trust in the context of online forums? How do the individual dimensions of trust contribute to overall trust development within online forums?* Study two seeks to add knowledge into literature reviews by investigating the development of trust within online forums.

Theories of critical mass have been applied in the field of economy (Marwell *et al.*, 1988), social psychology (Cho, 2011), information system (Westland, 2010) and marketing (Wattal *et al.*, 2010). The concept of critical mass discusses the processes from local to mass movements (Oliver *et al.*, 1993), and has dual natures. On one hand, it involves phase transitions in term of the growth rate of the numbers of participants that evolves from the local to mass movements. Phase transition is understood through the percolation theory applied to the network structural analysis (e.g. Centola, 2013). On the other hand, it considers the social influence and normative pressure on individuals' behaviours (Marwell *et al.*, 1988; Roger and Todd, 1995). However, it is difficult to measure the critical point defined by the critical mass theory with empirical studies (Shen *et al.*, 2013). Previous studies of the theory of critical mass mostly focus on phenomena after the critical point (Cho, 2011). However, *how can the theory of critical mass be applied to understand the structural influence of online forums in relation to knowledge contribution*

continuance? How is the critical point beyond which a mass phenomenon of knowledge sharing within online forums achieved? What happens before and after the critical point in terms of the online knowledge contributions? To fill this gap in knowledge, measurement of the critical point that involves the structural considerations is derived from complex network science and will be discussed in study three.

Online trust and critical mass are proposed as the key influential factors on sustaining online knowledge contributions. The following sections will further discuss the two concepts implied in the context of online forums.

2.3 Theoretical background

This section seeks to develop hypotheses and explore the dynamic aspects of the identified antecedents within the proposed integrative model. In section 2.3.1, it firstly conceptualizes the multi-dimensional construct online trust and its importance in the online context. Secondly, the concept of perceived critical mass that is relevant to the theory of critical mass is explained in section 2.3.2. Having discussed why the identified dynamic antecedents, i.e. online trust and perceived critical mass, play an essential role in online knowledge contribution behaviours, it thirdly proposes an integrative model that can predict the ongoing online knowledge contribution behaviours with section 2.3.3. Within this section, the causal relationships between these antecedents are further discussed. Finally, in order to investigate the dynamic aspects of antecedents, section 2.3.4 discusses the development of online trust, and section 2.3.5 explains the evolution of perceived critical mass using theories borrowed from network science.

2.3.1 Conceptualizing online trust

Human factors are important in the social computing area – even more so in the case of an online forum, which can only survive with the involvement of people. As discussed above, online information can be considered to be a public good that is associated with social dilemma and collective actions (Ostrom, 2000; Ostrom, 2010). Social dilemma in the context of online communities implies that members must make explicit or implicit assessments about the consequences – to themselves and others – arising from

the possibility of free-riding (Palmer and Huo, 2013). Collective actions that focus on voluntary contributions of members can overcome the problem of social dilemma. According to Ridings *et al.* (2002), trust is a fundamental characteristic associated with collective actions and the social dilemma.

Some researchers have argued that trust associated with human properties is not a construct that can be applied to computers, systems and online information gathering, which may explain why there is a relative dearth of trust studies in information science (Kelton *et al.*, 2008). However, trust that acts as an indicator giving an incentive to individuals' participation in knowledge sharing cannot be ignored (Chiu *et al.*, 2006).

Prior research identifies "reputation" (Han *et al.*, 2009), "reciprocity" (Chan and Li, 2008) and "commitment" (Kim *et al.*, 2008) as key elements on online social relationships. However, according to McKnight and Chervany (2002), the definition of "trust" should not be confused with "trust beliefs" and "the outcomes of trust". "Trust beliefs" are the antecedents of trust behaviours, while "the outcomes of trust" such as "reputation", "involvement" and "commitment" are the consequences of trust (McKnight and Chervany, 2002). This is beyond the research of this thesis, because it seeks to define trust in term of its levels and dimensionality within online forums, and to identify the role of trust on the online voluntary contribution behaviours, as discussed in the followings.

2.3.1.1 The levels of trust

There have been many conceptualizations of trust and its dimensionality. The simplest definitions of trust have conceptualized it as a disconfirmation state where expectations about an event, based on other parties' promises, are not matched by actual outcomes (Hosmer, 1995). Trust has been distinguished into three main categories (McKnight *et al.*, 1998): individual trust, interpersonal trust and impersonal institutional trust.

Individual and interpersonal trust have a similar definition, referring to an implicit set of beliefs that other(s) will not perform opportunist behaviors (Ridings *et al.*, 2002). Individual trust occurs between a

pair of individuals, while interpersonal trust is attributable to the relationships within or between the social groups such as communities (e.g. McKnight and Cheveny, 2002). In the contexts where individual(s) is/are dealing with an organization, the concept of institution-based trust, or institutional trust, has been advanced to incorporate trust in organizational mechanisms and structures (Hosmer, 1995; McKnight and Cheveny 2002). Institutional trust derives from a trustor's beliefs that favorable conditions, including ethical norms (situation normality), governance structures and technical capability (structural insurance) are present within an institution and provide safe conditions in which trustors make themselves vulnerable to possible non-fulfillment of the institution's promises (Pavlou, 2002).

2.3.1.2 The dimensionality of trust

Some researchers have seen trust as a uni-dimensional construct (Larzelere and Huston, 1980); however, a large body of literature has identified multiple dimensions of trust. For instance, the dimensionality of trust is a measure of belief in the benevolence and competence of others (e.g. Pavlou, 2002). There has been debate about whether trust and distrust are polar extremes of a single construct, or whether they are distinct constructs. Table 1 summarizes dimensions of trust which have been identified in selected studies, and the dominant disciplinary perspectives from which they are derived.

Table 1: The dimensions of trust

Dimensions of trust identified	Perspectives	Author(s)
Competence, Integrity, Predictability, Attraction	Sociology	Giffin, 1967
Benevolence	Psychology	Larzelere and Huston, 1980
Competence, Integrity	Psychology	Barber, 1983
Ability, Benevolence, Integrity	Sociology	Butler, 1999
Ability, Benevolence, Integrity, Predictability	Economics & Sociology	McKnight and Cheveny, 2002
Credibility, Benevolence	Economics	Pavlou, 2002
Ability, Benevolence	Economics	Ridings <i>et al.</i> , 2002
Ability, Benevolence,	Economics	Das & Teng, 2004
Competence, Benevolence	Sociology	Cho, 2006
Ability, Benevolence, Integrity, Predictability	Sociology & Economics	Brüttner and Göritz, 2008
Ability, Benevolence, Integrity	Sociology & Economics	Lu <i>et al.</i> , 2009
Ability, Benevolence, Integrity,	Sociology &	Wu <i>et al.</i> , 2010

Predictability	Economics	
Credibility, Benevolence	Psychology	Dimoka, 2010

Based on the above-mentioned trust dimensions, four themes can be identified: ability, benevolence, integrity and predictability. McKnight and Chervany (2002) study the relationships between consumers and Internet sellers. Embedded in their definitions of trust beliefs, the following understandings are derived. Ability is defined as skills or competencies that permit an individual to have a certain power in a specific area or domain. Benevolence refers to the expectation that trustees, who are caring and intend to help, would like to give appropriate advice and support to trustors. Benevolence is the most strongly linked with trustees' attitude concerning affection and confidence. Integrity means that the expectation that trustees will act in accordance with what is expected by trustors, such as not telling a lie, or acting dishonestly. Integrity is often revealed through trustees' behaviours. Predictability is associated with one's secure feelings of others' behaviours that are constantly predictable (the prediction of others' behaviours can be both good and bad).

The combinations of the four dimensionalities of trust are applied in this thesis to understand the different levels of trust in the context of online knowledge contributions. In the context of online forums, this corresponds to trust in individual members, and trust in the forum as a collective entirety. Trust in members is associated with individual and interpersonal trust. Members may trust in one particular member who can provide valuable ideas, but in most cases they communicate with the collective entireties. In contrast, trust in online forums is understood as institutional trust, because members may regard an online forum as an organisation or institution and evaluate its mechanisms and conditions. From this evidence, the study seeks to extend knowledge by considering the case where both trust in members and trust in online forums impact on the ongoing intention of knowledge sharing within the forum.

Most research has tended to present a partial perspective of trust but does not fully capture its complexity.

Although online trust has been articulated as comprising interpersonal trust (Wasco and Faraj, 2005) and

institutional trust (Hsu *et al.*, 2007; Zimmer *et al.*, 2010), these different levels of online trust interact and their influence on continuing online knowledge contribution behaviour remains under-researched.

2.3.2 Theory of critical mass and perceived critical pass

This section seeks to explain how theory of critical mass can be implied in the context of online forums. First, it explains why the theory of critical mass is relevant in order to understand the online voluntary contribution behaviours. Second, the theory of critical mass and the underlying assumptions of the theory are discussed. Third, the concept of perceived critical mass that is part of theory of critical mass is explained.

2.3.2.1 The importance of theory of critical mass

Online knowledge is characteristic of digital public goods (Wasco *et al.*, 2009), and overcoming the social dilemma problem relating to public goods requires collective actions that are specifically embedded in voluntary cooperation (Ostrom, 2000; Wasco *et al.*, 2009). Because public goods are characterized of non-excludability and non-rivalry, an individual can benefit from selfishness if there is enough voluntary contribution to public goods. In the case that everyone chooses the selfish alternative, the whole group loses (e.g. Ostrom, 2000). Explanations for why people contribute to online forums have been derived from economic, sociological and biological perspectives.

One significant contribution derives from game theory, which has been refined with models of direct and indirect reciprocity to incorporate an individual's expected return from their contribution (Axelrod and Hamilton, 1981; Imhof *et al.*, 2007; Taylor and Nowak, 2007; Nowak *et al.*, 2010). Theories of altruism have sought to explain why an individual takes care directly of others' welfare at their own cost (Frohlich, 1974), with expected returns for altruistic activities, which is proposed by genetic kinship theories (Wright, 1922; Hamilton, 1964; Dawkins, 1976; Alger, 2010). Hamilton (1964) argues that animals such as ants and honey bees help their kin because their altruistic behaviours can increase their inclusive fitness.

However, there are different underpinning assumptions for these theories. Altruism behaviours often present in homogenous populations, while reciprocity requires conditions of simultaneous contribution (Nowak *et al.*, 2010). Wasco *et al.* (2009) describe online forums as types of e-communities that are made up of a dynamic-continuous inflow and outflow of heterogeneous members. This calls for theories that can explain the multiple processes rather than the simultaneous contribution. That is, traditional theories of altruism and reciprocity have not taken sufficient account of the scale of online networks relative to their heterogeneous population and subsequent activities in influencing contribution to a forum. Wasco *et al.* (2009) argue that the theory of critical mass (Oliver and Marwell, 1988) is relevant for understanding the creation of digital public goods, because it overcomes restraint assumptions associated with both theories of altruism as well as reciprocity.

Network theory suggests that the value of a network grows with the increasing numbers of members (Li *et al.*, 2011). The value of an interactive communication system for an individual member is determined by the numbers of users of that system (Cho, 2011). An online forum provides an opportunity for members to exchange ideas with a large number of others (Li, 2011), suggesting that the intrinsic value of online forums increases with the expansion of participation and knowledge sharing. The mass phenomenon can be understood through the theory of critical mass (Oliver *et al.*, 1985; Roger and Todd, 1995; Wasco *et al.*, 2009; Centola, 2013).

Critical mass theory (Oliver *et al.*, 1985; Marwell *et al.*, 1988; Oliver and Marwell, 1988; Pahl *et al.*, 1991) incorporates theories of individuals' rational choice into collective action (Centola, 2013) and argues that a group of initial contributors can pay the set-up cost and thereafter promote the mass contribution beyond a critical point. Critical mass theory (Oliver *et al.*, 1985) explains how a small number of selected individuals can have a powerful, positive impact on mass collective production. Similar to threshold models (Granovetter, 1978); it focuses on the number or proportion of initial self-interested contributors for whom net benefit exceeds net cost. This phase transition is analysed through the contagion model in biology (Dodds and Watts, 2004), and self-organised criticality in physics (Bak *et*

al., 1987). In social life, one simple example is “fashion”, where several selected celebrities can evoke uniform massive behaviours. Online forums involve a minimum number of initial contributors. The following section examines the concept of initial contributors in a large-scale network contextualized in the theory of critical mass.

2.3.2.2 The importance of initial contributors in theory of critical mass

Critical mass theory is the most compelling argument of Olson’s (1965) logic of collective action (Oliver and Marwell, 2001). Olson (1965) points out that a rational individual will not behave cooperatively in order to achieve a common or general interest, without incentive or punishment mechanisms that reward co-operators or punish non-co-operators. Marwell and Oliver (1993) argue that initial contributors can create positive incentives for subsequent actors, which generates a widespread influence over the group to support the production of public goods. According to Marwell and Oliver (1993), there is a possible self-reinforcing system in collective actions and it is the initial contributors who pay the set-up cost and promote future contribution behaviours of subsequence actors.

The original critical mass model developed by Marwell *et al.* (1988) can be employed to illustrate individuals’ decisions about contributing to public goods as follows:

$$G = p(\sum r)I - r, \quad (2.1)$$

where G represents an individual’s net gain from contribution. It interprets the relationship between an individual and the group in general. Thus, it omits the interactions between individuals but highlights the general exchange pattern. $p(\sum r)$ refers to the production function of the total contribution by all parties to public goods, which specifies the relationship between inputs of total resource contribution and outputs of levels of public goods. Furthermore, the production function in this model is a u-concave or accelerating function, which facilitates increasing marginal returns. In online discussion, for instance, one response to a seed message may tell one piece of the “truth”, the second one contributes to another piece. In other words, an accelerating production function encourages individuals to make sequential contributions that

are embedded in previous outputs, because additional contributions could accelerate achievement of certainty. However, the central challenge is to start collective actions because rational individuals will contribute in the late stage in order to enjoy higher payoffs. I is an individual's interest level in the public good, and r means an individual's contribution resource. That is, when $p(\sum r) > r/I$, the total payoff from all contributions to public goods exceeds the individual's r/I ratio, and an individual will make a positive contribution decision. In other words, the value of a given public good is subjected to available resources and the willingness to pay: the higher the interest level, the more possible that an individual contributes. The richer the resources are available, the bigger the outputs can be.

It can be concluded that there are two important assumptions in the critical mass model: the accelerating production function that highlights the feasibility problem, and the group heterogeneity. Feasibility refers to finding the initial point that satisfies all constraints in the u-concave production function. The group heterogeneity allows either highly interested or resourceful individuals to pay the early start-up cost of collective actions. The idea of critical mass is related to exactly these kinds of contributors. In this sense, the critical mass members attract numerous others to contribute sequentially. The emergence of the initial contributors is therefore highlighted in solving the feasibility problem (Centola, 2013). It is noted that the initial contributors are not necessarily the critical mass contributors emerged in a later stage of development of online networks.

2.3.2.3 Perceived critical mass

The literature cited above indicates that critical mass contribution models describe a transition model that leads to a stable and dynamic collaboration (Oliver *et al.*, 1985), and that it can be applied to understand members' voluntary contribution behaviour in online forums (Wasco *et al.*, 2009). The studies of the theory of critical mass have been divided into understanding the concept of perceived critical mass (Roger, 1995; Cho, 2011; Shen *et al.*, 2013) and the formal matrix of critical mass (Westland, 2010; Centola, 2013). Perceived critical mass describes the mass phenomenon after the critical point, and it is understood from a social influence point of view (Cho, 2011; Shen *et al.*, 2013). The formal matrix of

critical mass tends to explore what has happened before and after the critical point, which involves of the network analyses (Centola, 2013). The notion of critical point is associated with the phase transition (Oliver *et al.*, 1985). An example of such is the changes in opposite direction along a curve. To use another example, water becomes steam after the critical point. In the context of online forums, it refers to changes occurred after the critical point in terms of the growth rate.

Perceived critical mass in online forums refers to the perceived numbers of members within a group who have a social influence on knowledge contribution (Cho, 2011; Shen *et al.*, 2013). However, it is not clear how perceived critical mass interacts with the different levels of online trust and the role it plays in online knowledge contribution continuance. For instance, the perceived mass numbers of knowledge contributors within an internet-based communication system can accelerate future collective knowledge sharing (Shen *et al.*, 2013). Yet, the important role of perceived critical mass has been examined in a nested model with respect to intentional knowledge contribution by neglecting other important variables (Shen *et al.*, 2013). More empirical evidence is required of the dynamic antecedents of online knowledge contribution.

The formal matrix of critical mass has been studied with simulations within random networks (Centola, 2013). However, the functionalities of networks vary with the type of network (Newman, 2003). The study of Barabasi and Albert (1999) shows that many real world networks are scale-free with variation between neighbours associated to a member that tends to be infinite. Many web-based communities are set within this category (Newman, 2005). Against to this background, this thesis seeks to extend knowledge of the theory of critical mass applied within online forums, with discussions presented in section 2.5.

With respect to online forums, the above discussions can be summarized in six principles: 1) the essential reason why the critical mass members contribute relies on their perception that benefits of contribution exceed the contribution cost; 2) the contribution by a small number of critical mass members will balance the ratio of benefit/cost of the whole network. (although the problem of “free riders” persists in an online

forum, there is evidence that the digital public goods can always be created by a small group of contributors (Wasco *et al.*, 2009)); 3) the more benefit from contribution, the more likely that the critical mass members are active; 4) the size of online forums is often associated with the possibility of the emergence of the critical mass members, because the bigger the size, it is more likely to have an increasing marginal return for contributions. (Wasco *et al.* (2009) argue that the size of online discussion groups is associated with the value of that group); 5) the critical mass members often play the role of “bridge” between the communications of members, because it is critical mass members who participate in the majority of discussions within online forums; 6) there is a critical point after which a mass phenomenon is observed.

The above has explained the complex concepts of the key antecedents identified in this thesis. The following section 2.3.3 will discuss how these antecedents act together and impact on the continuous online knowledge sharing.

2.3.3 Predicting continuous online knowledge sharing behaviour

This section seeks to develop hypotheses that identify the causal relationships within the proposed integrative model. The conceptual model embedded in DTPB is firstly presented. Each construct and relationship presented in the model is explained thereafter.

2.3.3.1 Introduction of the theoretic background

The conceptual development is drawn on theory of planned behaviour (TPB) (Ajzen, 1991) from social science and decomposed theory of planned behaviour (DTPB) (Taylor and Todd, 1995), derived from TPB.

Ajzen (1991) argues that behaviour is determined by intention which can be measured through the attitude towards the behaviour, the subjective norms regarding the behaviour, and the perceived behavioural control (PBC) about one's behaviour. Attitude is based on a set of beliefs in the expectation of behavioural consequences which may be either positive or negative. Subjective norms describe one's

perceived social pressure with regard to engage or not to engage in the behaviour. PBC is evaluated through one's control beliefs in abilities, reflected in resources or skills, and the opportunity of engaging in the behaviour (Gagné, 2009).

One significant contribution of TPB is its articulation of the causal relationships between intention and behaviour. Numerous empirical studies in the field of knowledge sharing have proposed intention as an aggregation of motivational factors that express individuals' willingness to perform the behaviour in question, and it is an immediate antecedent of behaviour (Ajzen, 1991; Szulanski, 1996; Szulanski and Jensen, 2006; Erden *et al.*, 2012). The entire three components are measured by aggregating beliefs that are the summed evaluation of personal experiences over time (Ajzen, 2011). As a general rule, the stronger the intention, the more likely an individual is to perform behaviour. Therefore, the underlying assumption of this study is that intention will lead to behaviour.

Although TPB claims to be a generalized model that has been widely applied, specific antecedents vary according to the context (Taylor and Todd, 1995a). To overcome this limitation, DTPB decomposes the antecedent components in TPB in order to provide greater understanding of the intentional behaviour in a particular context (Taylor and Todd, 1995). Following the TPB logic, DTPB expands the scope of the model by considering the fact that behavioural antecedents vary according to the context (Taylor and Todd, 1995a). DTPB decomposes the existing antecedent components in TPB, providing greater understanding of the intentional behaviour in a particular context (Taylor and Todd, 1995) and the understanding of its determinants (Chennamaneni *et al.*, 2012). For instance, DTPB can be developed for understanding knowledge sharing within organisations by decomposing the attitudinal beliefs into perceived organisational incentives, perceived reputation enhancement and perceived enjoyment in helping others (Chennamaneni *et al.*, 2012). The emphasis of DTPB is not to test the immediate determinants of intention, or the logic of TPB, but to explore the contextual factors of these determinants in order to provide a complete comprehension on the intentional behaviour in a particular context (Taylor

and Todd 1995). DTPB is introduced in this study to allow behavioural, normative and control beliefs to be decomposed into multiple dimensions to fit a specific context (Rogers, 1995), i.e. online forums.

Table 2: *Determinants of TPB*

Belief toward an outcome Evaluation of the outcome	Attitude	Intention	Behaviour
Beliefs of what others think, what experts think, and motivation to comply with others	Subjective norm		
Beliefs of self-controllability with respect to skills, resource and opportunity of performing a behaviour.	Perceived behavioural control		

Adopted from Ajzen (1991)

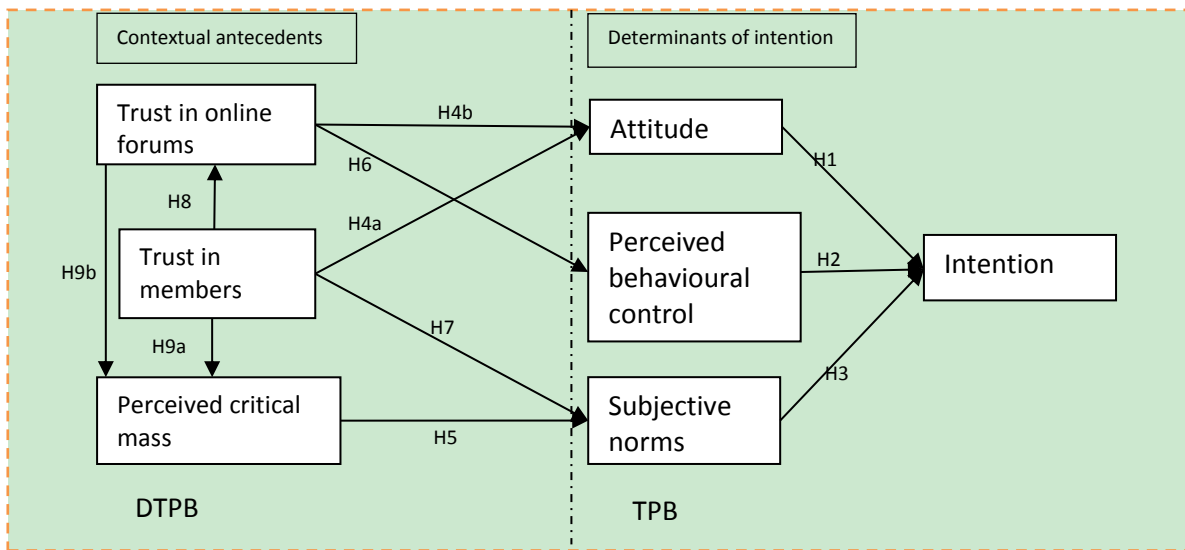
Decomposed TPB (DTPB) is introduced where the behavioural, normative and control beliefs are decomposed into multiple dimensions to fit the situation in question (Rogers, 1995; Chennamanenia *et al.*, 2012; Sahli and Legoh  rel, 2015), such as the usage of IT technology (Talyor and Todd, 1995), knowledge sharing (Chennamaneni *et al.*, 2012) and online shopping (Lin, 2007).

Previous research on knowledge sharing is mainly discussed in the context of intra-organisation studies (Finkbeiner, 2013). However, in the case of online forums, beliefs underlined knowledge contributors can be different from employees. For instance, organisations often use incentive or punishable instruments to encourage employees to participate knowledge sharing within them (Ba, 2001), which are not good solutions to online forums that require voluntary contributions.

In the context of online forums, continuing participation through ongoing discussions and willingness to share information play an important role in sustaining online forums (Harris and Rae, 2009). One of the main objectives of this study is to explore specific factors that have influence on ongoing online contribution intentional behaviour, and which facilitate the sustainability of online forums.

The conceptual framework is shown in Figure 3, and in the following paragraphs the relationships between the components that form the bases for the hypotheses will be discussed.

Figure 3: Conceptual model



2.3.3.2 Intention toward knowledge sharing within online forums

Knowledge provided through online forums can be considered a form of *public goods*, marked by its non-excludability of consumption (Wasco *et al.*, 2009). Public goods are outputs of collective contribution, and all individuals are able to access public goods regardless of their own personal contributions (Snidal 1979). Online knowledge contribution behaviours are typically embedded in members' voluntary intention to contribute, because online knowledge is characteristic of public goods that are neither excludable nor create rivalry in consumption (Wasco *et al.*, 2009).

One problem associated with public goods is social dilemma (Ostrom 2010) in which knowledge sharing may be considered at the individual and group levels (Lin, 2007). Individually-based decision making models indicate optimisation of an individual's preference function subject to informational and material constraints (Gintis, 2007). Under this logic, when every individual is rational and enjoys a public good for free, the public good will never be produced. Members of an online forum may benefit from the knowledge provided by forum members but make little or no contribution – these are often labeled as “free riders” (Wasco *et al.*, 2009). Overcoming the public goods problem requires collective actions that are specifically embedded in voluntary cooperative interactions between members (Ostrom, 2000).

Online knowledge contributions are cooperative and emergent processes. Members make sequential contributions that are embedded in previous outputs. For example, one response to a seed message may contribute one piece of knowledge; a second response may contribute another part and so forth. In this sense, knowledge provided through online forums can have value-added (Wasco *et al.*, 2009).

Following the idea that collective actions are a solution to the public goods dilemma (Ostrom, 2000), it is important to understand online knowledge sharing behaviours, not only on the individual but also the collective levels. Indeed, knowledge sharing has been studied at the individual and group levels (Lin, 2007). There is a consideration relating to the distinguishable notion between an individual's intention and "we-intention". Individual's intention is associated with the motivational aggregated factors that have an effect on the intended behaviour (Ajzen, 1991). "We-intention" reflects the social influences on members' behaviours within virtual communities (Bagozzi and Dholakia, 2006; Shen *et al.*, 2009). In these conditions, an individual regards him/herself as a group member and performs collective actions with others (Bagozzi and Dholakia, 2002). This assumes that one will adopt one's intentional behaviours in order to achieve a common goal that would satisfy all members within the group (Shen *et al.*, 2013). Online knowledge sharing represents a cooperative process among members who seek knowledge provided by other members (Wasco *et al.*, 2009). As a result of which, it is considered that the intention to contribute knowledge within online forums is also embedded in the collective levels. This understanding provides the base for the different levels of trust studied in the context of online forums, further explained in the section 2.5.4.1.

Additionally, philosophers have distinguished intention as a present-oriented intention and future-oriented intention (Cohen and Levesque, 1990). The present-oriented intention refers to what will be done in a short time while the future-oriented intention is a stronger willingness to perform actions that involves planning and adoption of behaviour (Bratman, 1984). It is argued that the future-oriented intention that involves commitment to plan to behave (Bratman, 1984) could be a good reflection of behaviour pattern. As a summary of the above discussions, intention is understood as the "future-oriented intention"

demonstrated by members rather than a member of an online forum. Thereby, this study seeks to identify the dynamic antecedents of members' intentions to engage in online knowledge sharing.

2.3.3.3 The determinants of intention

TPB (Ajzen, 1991) identifies behavioural intention (BI) as a function of beliefs that are influenced by attitude (A), perceived behavioural control (PBC) and subjective norms (SN). It is worth mentioning that, the objective is not to test the logic of TPB that has been extremely researched (e.g. Erden *et al.*, 2012), rather, it seeks to decompose the determinants of TPB in order to provide a precise and richer understanding of TPB applied in the online context. The following will further explain the three determinants of behavioural intention.

Attitude to online knowledge contribution

Attitude (A) is the sum of attitudinal beliefs (b_i) in a particular consequence of performing behaviour, which is weighted by the evaluation of the desirability of that expected result (e_i) (Ajzen, 1991; Talyor and Todd, 1995). It can be expressed as: $A = \sum b_i e_i$ (2.2).

Empirical studies have shown that attitudinal beliefs associated with the expected outcomes positively influence intentional behaviours (Ajzen, 1991). An example raised by Talyor and Todd (1995) explains that one may believe that using IT will enhance job performance, and this belief is evaluated as a highly desirable outcome. In the context of discussions within online forums, an individual may believe exchanging ideas with others will be helpful (b_i), because discussions added to a topic left by members are often value-added (Wasco *et al.*, 2009), and this belief is evaluated as a highly desirable outcome (e_i). The following hypothesis is then derived:

H1 Attitudinal beliefs have a positive influence on intention to the continuous sharing of knowledge online.

PBC beliefs of knowledge sharing within online forums

Perceived behavioural control in the continuous use of new technology has been characterised by two dimensions, which are controllability and the self-efficiency of adaption of learning new technology (Taylor and Todd, 1995; Ajzen, 2002a). Self-efficiency refers to users' confidence in their ability for using technology (Taylor and Todd, 1995) and is further enhanced when the technology is easy to use and users feel a sense of control when using it (Venkatesh and Goyal, 2010). Perceived behavioural control is the sum of the set of control beliefs (cb_k) which is weighted by the perceived facilitations (pf_k), and can be expressed as: $PBC = \sum cb_k pf_k$ (2.3) (Ajzen, 1991; Tylor *et al.*, 1995).

Perceived behavioural control has been shown to be positively associated with intention to share knowledge within online communities (Chennamanenia *et al.*, 2012; Erden *et al.*, 2012; Finkbeiner, 2013). In the case of online knowledge sharing, an individual might be confident in his/ her ability to contribute knowledge (self-efficiency), and acknowledge the support provided by an online forum (controllability) which can for example allow members to enjoy participating in its knowledge sharing activities. The following hypothesis is therefore derived:

H2 PBC has a positive impact on members' intention to continuously share knowledge.

Subjective norm beliefs of knowledge sharing within online forums

Subjective norm is defined as the sum of individual's normative beliefs (nb_j) of significant others' willingness on his/her part to perform a particular behaviour, which is weighted by that individual's motivation to comply others' willingness (mc_j) (Ajzen, 1991; Tylor, 1995). The subjective norm is expressed as: $SN = \sum nb_j mc_j$ (2.4) (Ajzen, 1991; Tylor, 1995).

With respect to online knowledge sharing, subjective norms can be understood as one's perception of normative pressure on cooperative knowledge contribution and one's motivation to respond to such pressure (Jeffries and Becker, 2008). It is because cooperation is essential to solve public goods dilemmas

(Ostrom, 2000), including the actions of online communities that require cooperation among members (Yen *et al.*, 2011); subjective norms within this context may be associated with norms of cooperation (Yen *et al.*, 2011). The norms of cooperation help members of an online community to understand what is expected from others and what to do in a given context, which are often inferred from text-based communications (Yen *et al.*, 2011). In this context, normative beliefs refer to individual or collective beliefs about the prescribed/proscribed behaviours under a particular social context (Bicchieri and Muldoon, 2014). If members perceive an online community to be generally supportive of cooperation, they are more likely to integrate cooperative norms as their behavioural guidelines (Yen *et al.*, 2011), as a result of which, they are more likely to participate in online discussions. Hence, it is hypothesized:

H3 Subjective norm is positively associated with intention to share knowledge online.

2.3.3.4 Contextual factors for online contribution

TPB has been criticised, among other reasons, because of its inability to incorporate contextual factors in the decision making process. In the following paragraphs two factors are reviewed – trust and perceived critical mass, which the literature has suggested, may be important antecedents of online contribution intention. These variables are therefore integrated within the conceptual framework of DTPB that seeks to decompose the beliefs of attitude, perceived behavioural control and subjective norms.

Online trust

Most research has tended to present a partial perspective of trust, not fully capturing its complexity. In the context of online forums, interpersonal and institutional trust corresponds to trust in individual members, and trust in the forum as an entity, respectively. Members may trust one particular member who can provide valuable knowledge, but in most cases they communicate with collective entities. Trust in online forums is understood as institutional trust, because members may regard an online forum as an organisation or institution and evaluate its mechanisms and conditions that will favour knowledge sharing with members. Although online trust has been articulated as comprising interpersonal trust (Wasco and Faraj, 2005) and institutional trust (Hsu *et al.*, 2007; Zimmer *et al.*, 2010), these different types of online

trust interact and their influence on continuing online knowledge contribution behaviour remains under-researched. The study extends existing theory by considering the case where both trust in members and trust in online forums impact on the ongoing attitudinal intention of knowledge sharing within the forum.

Trust can be linked to social recognition (e.g. McKnight and Chervany, 2002). Social recognition generally involves social achievements vis-à-vis others that one deserves (Brandom, 1994). With respect to online knowledge contributions, it can refer to, for example, members' ability to share knowledge is publicly acknowledged by others. Social recognition can be both an antecedent and consequence of interpersonal trust (McKnight and Chervany, 2002). Research on knowledge and resource management within communities suggests that evaluations of desired outcomes such as social recognition will positively influence attitudinal beliefs to share knowledge (Jiang *et al.* 2002). It is therefore argued that trust in members impacts on the attitudinal knowledge sharing. In addition, the mechanisms and competences of a community in handling communications between members contribute to institutional trust (McKnight *et al.*, 1998). The two dimensionalities of institutional trust, i.e. situational normality and structure insurance, have an effect on individuals' ongoing online knowledge sharing attitudinal behaviours (Chen, 2007). Situation normality describes a situation within which an event or a phenomenon is explicable, and structural insurance involves ensuring contextual structures and technology that will facilitate the expected success (McKnight and Cheveny, 2002). Thus, the following hypotheses h4a and h4b are derived embedded in the above discussions are proposed:

H4a Trust in members of an online forum positively influences attitude to knowledge sharing.

H4b Trust in an online forum positively influences attitude to knowledge sharing.

Perceived behavioural control beliefs are influenced by perceived facilitative conditions (Ajzen, 1991; Talor and Todd, 1995a). Online communities provide a platform and opportunities for members to explore their knowledge and capacities (Erden *et al.*, 2012; Wasco *et al.*, 2009), and contributors are more likely to share knowledge if they feel safe as a result of believing they have control over their behaviour (Erden *et al.*, 2012). For instance, when members process their private information through an online

channel, they may feel anxious about the willingness of recipients to treat such information confidentially. As another example, members may also reduce their online contribution when the time taken to upload and download forum pages is perceived as excessive (Wasco *et al.*, 2009). It is suggested therefore that perceived trust in the competence and facilitative conditions of an online forum (for example with regard to its resources and technology) will influence members' PBC. The following hypothesis is derived:

H5 Trust in online forums has a positive impact on perceived behavioural control.

Subjective norms in the context of text-based communications is particularly associated with norms of cooperation (Yen *et al.*, 2011), which is underpinned by interpersonal trust (Coleman, 1988). A trust belief that working partners will not behave opportunistically leads to long-term relationship investment (Morgan and Hunt, 1994) and information sharing (Dyer and Chu, 2003). Jeffries and Becker (2008) argue that the relationship between trust in others within an organisation, and the intention to be cooperative occur indirectly through the perception of subjective cooperative norms. Intention to cooperate can exist when the subjective cooperative norms are supportive, even though employees have low levels of trust in others (Jeffries and Becker, 2008). However, high levels of trust in others may lead to social influences on intended behaviours (Jeffries and Becker, 2008). For instance, an employee may perceive that his/her peers find the use of new technology to be important. If the employee agrees with this thought, he/she is more likely to adopt the new technology (Tylor and Todd, 1995). Similarly, it is argued that the more members believe that their fellow members are cooperative, the more they are likely to acknowledge the normative pressure on their own intended behaviours. It is therefore argued that trust in members is an antecedent of subjective norms belief, which leads to the following hypothesis:

H6 Trust in members positively influences subjective norms.

Perceived critical mass

Critical mass theory (Oliver *et al.*, 1985; Marwell *et al.*, 1988; Oliver and Marwell, 1988; Prahl *et al.*, 1991) entails insights about individuals' rational choice into collective actions (Centola, 2013). The

theory argues that a group of initial contributors can pay the starting costs and thereafter promote the mass contribution within a collective group. Marwell and Oliver (1988) argue that there is a possible self-reinforcing system in collective actions, and the system is sustained by a small group of members. This argument is based on two important assumptions (Marwell and Oliver, 1993): (i) group heterogeneity where a small group of members have more resources to contribute; (ii) an accelerating production function making the outputs of collective actions exceed what could be achieved by an individual (hence attracting more followers). After a critical point, a mass phenomenon occurs. Perceived critical mass that describes the mass phenomenon after the critical point is part of theory of critical mass. Due to the difficulty of measuring the critical point that has been highlighted in the theory of critical mass, the assessment of critical mass is been replaced by measuring the perceived size of usage of a system / new technology (Cho, 2011). Section 2.3.5 seeks to address this limitation.

Perceived critical mass in online forums refers to perceived numbers of members within a group that have a social influence on knowledge contribution (Cho, 2011; Shen *et al.*, 2013). Limited studies have sought to examine the role that perceived critical mass plays in online knowledge contribution. Shen *et al.* (2013) argue that perceived critical mass can accelerate future collective knowledge sharing within an internet-based communication system. However, this argument has been examined in a nested model by neglecting other important variables such as trust and perceived behavioural control (Shen *et al.*, 2013). As a consequence, more empirical evidence is required to understand why the notion of perceived critical mass may contribute to the dynamic and inter-related antecedents of online knowledge contributions.

Previous studies have found that normative belief is associated with perceived critical mass (Cho, 2011; Shen *et al.*, 2013). Normative beliefs refer to individual or collective beliefs about the prescribed/proscribed behaviours under a particular social context (Bicchieri and Muldoon, 2014). Though normative beliefs alone cannot support the subjective norms that are also weighted by the motivations to perform a particular behaviour (e.g. Ajzen, 1990), they are essential to understand subjective perceived norms that have been widely recognised as being efficient to solve problems in

question (e.g. Posner, 2000). The concept of perceived critical mass considers the social influence and normative pressure on individuals' behaviours (Roger and Todd, 1995; Cho, 2011; Shen *et al.*, 2013). Studies on innovation diffusion have shown that that one's subjective perception of critical mass usage of innovation will evoke social influences and give the normative pressure on the subsequent adaption by oneself (Roger and Todd, 1995; Wattal *et al.*, 2010). With respect to the usage of online forums, it is argued that members' subjective perception of the existence of a critical mass number of members who contribute knowledge online, can give them the normative pressure to be cooperative, because members may like to be the seen as 'normal' within an online forum (Cho, 2011). Hence, the following hypothesis is derived:

H7 Perceived critical mass has a positive influence on subjective norm.

2.3.3.5 The causal relationships between antecedents

Both interpersonal and institutional trusts involve multiple dyadic relationships (Ferrin and Dirks, 2006). Following cumulative successful social exchanges, an individual may increase her/his expectation of and confidence in the return of goodwill by others (Luo, 2006). A member may benefit from the suggestions provided by another member and increase his/her confidence in members in general who can provide valuable knowledge (Luo, 2006). In other words, interpersonal trust can facilitate the development of impersonal institutional trust (Luo, 2006). There is previous evidence in the context of online auction websites that trust in other users can lead to general trust in an auction website (Schlosser *et al.*, 2006). Similarly, trust in members who will voluntarily contribute knowledge can lead to trust an online forum that is a good platform for sharing knowledge. Members may trust in an online forum because it can create an atmosphere marked by active knowledge sharing within members. Hence, the following hypothesis is derived:

H8 Trust in online forum members has a positive influence on trust in the online forum.

Within online social networks, a relationship develops between pairs of members if they have contact with each other, such as replying to an enquiry (Haythornthwaite, 2002). Interpersonal ties are often used in social science to describe the information-carrying connections between individuals (e.g. Haythornthwaite, 2002). This involves the strong or weak ties (Granovetter, 1973).

Frequent reciprocal contact between individuals is characteristic of strong ties (Granovetter, 1973). The strength of a tie is a combination of the amount of time, mutual confiding and the reciprocal services that one provides to another (Homan, 1950). Moreover, strong ties often occur between individuals who have common interests and are similar to each other in a various way (Granvovetter, 1973).

A high level of homophile that is associated with trust among individuals has been found to be effective in solving start-up problem in relationships. Empirical studies of computer-mediated communication have found out that repeated communications can lead to the development of trust (Haythornthwaite, 2002). In online networks, strong ties of this nature may facilitate the emergence of a critical mass of members (Centola, 2013). Because perceived critical mass only occur if there is a critical mass of members, it is argued that interpersonal trust, i.e. trust in members, can facilitate the perception of a large size of contributing members. Therefore, the following hypothesis is derived:

H9a Trust in members has a positive influence on perceived critical mass.

Haythornthwaite (2002) proposes “weak ties” to describe interactions with diverse peers within a broad communication system. Contrary to strong ties, weak ties don’t combine the mutual devoted amount of time and the reciprocal interactions between a pair. Weak ties can suggest that the probability of two untied parties becoming tied will increase when these two parties have common acquaintance(s) (Granovetter, 1973). Haythornthwaite (2002) argues that individuals are weakly tied because they have a sense of belonging to an institution. The sense of belonging is often used to measure the commitment to an event or institutions (e.g. Turri *et al.*, 2013). Commitment to an organisation is an antecedent and consequence of institutional trust (McKnight and Chevery, 2002). Given this understanding, trust in

online forums involves ‘weak’ ties among members who have heterogeneous resources to contribute online.

Previous studies have suggested that weak ties can span numerous interactions between peers (Granovetter, 1973) and promote membership collision (Centola, 2013). In the case of online forums, a trusted online forum is more likely to host millions of members than an online forum in reputeless. This leads to the understanding that the institutional trust can positively impact on the mass usages of an online forum. Therefore, it is argued that trust in online forums can be another sufficient condition for the existence of the perceived critical mass:

H9b Trust in online forums has a positive influence on perceived critical mass.

Baek and Jung (2015) argue that institutional trust is the key mediator of the relationships between interpersonal trust and the commitment to organisations. They further argue that institutional trust can mediate the effects of interpersonal trust within organisations when they both occur. Because commitment is either the consequence or antecedent of general trust (Mcknight and Chervany, 2002), it implies that institutional trust mediates the effects of interpersonal trust within organisations. Given this understanding, it is plausible that trust in online forums (institutional trust) could mediate the influences of trust in members on perceived critical mass. Therefore, the following hypothesis is derived:

H9c Trust in online forums mediates the relationship between trust in members and perceived critical mass.

Study one has sought to investigate the effects of the key antecedents on the intentional contribution behaviours and the causal relationship between them. One major limitation of study one is that the dynamic nature of the identified antecedents is investigated within a cross-sectional design. Study two and study three are further developed to address this limitation, and will be discussed in sections 2.3.4 and 2.3.5 accordingly.

2. 3.4 The development of trust

The development of trust involves a gradual process to build (Li *et al.*, 2008) and to undermine (Charki and Josserand, 2008).

Through knowledge obtained during interaction, an individual may predict others' likely behaviours (McKnight and Chervany, 2002; Li *et al.*, 2008; Ermisch *et al.*, 2009). This process of trust building has variously been referred to as familiarity, knowledge-based trust, relational trust (Ferrin and Dirks, 2006) and process-based trust (Zucker, 1986; Grayson *et al.*, 2008). When the trustor is not yet directly familiar with the trustee, the concept of trust propensity has been used to describe the initial trust that the trustor has in the trustee. (McKnight and Chervany, 2002; Li *et al.*, 2008). Trust propensity, or disposition to trust, partly reflects an individual's personality and tendency to believe or not believe in others in general. For example, when information about a potential trustee is not available, an individual with strong faith in humanity might presume the other party to be trustworthy, whereas one with lower faith would be more likely to consult independent sources to assess likely trustworthiness (Deutsch, 1960a; Li *et al.*, 2008). In this context, referral through electronic media has become very important. Initial trust development is facilitated by the existence of networks within which information can be shared by actual and potential trustors. The perceived risk of mistrust is lowered where a potential trustor can use the accumulated knowledge of individuals who have previous experience of trusting an individual or organization (Zucker, 1986; McKnight and Chervany, 2002; Ermisch *et al.*, 2009; Li *et al.*, 2008; Lu *et al.*, 2009).

Web 2.0 technologies can greatly increase the ability of potential trustors to assess the trustworthiness of a potential partner, and many examples have been cited of brands that have been developed or undermined rapidly on the basis of comments about the trustworthiness of an organization spread virally using web 2.0 media. (de Valck *et al.*, 2009). Social network structure theory has been used to understand the nature of trust within virtual communities (Gilsing and Duysters, 2008; Ganley and Lampe, 2009; Wasco *et al.*, 2009), suggestion that the effect of trust assessment through online sources is dependent upon the structure and size of online communities, and the connectedness among members. The credence

given to information on which trust assessments are made is dependent on the strength of links between members of an online community and the degree of perceived similarity with contributors (Brown *et al.*, 2007). In games of trust, it has been noted that only trustworthy players can survive in the long-term where information about all players is complete (Ostrom, 2000).

Trust can be undermined over time, and the processes by which this occurs have been debated. Traditionally, trust and distrust have been considered as bipolar extremes of a single construct (Deutsch, 1958; Barber, 1983). Consequently, the absence of trust is regarded as a condition of distrust or low trust. However, Dimoka (2010) argues that distrust is founded on emotional deception rather than being a simple cognitive phenomenon that is present during the developmental stages of trust. Dimoka (2010) emphasized a multi-structure approach by noting that trust building is associated with the brain's cognition areas, while distrust is more strongly linked with the brain's processing of emotions, particularly relating to deception, fear of loss and unethical behaviour. Cho (2006) notes that trust reflects an individual's risk preference, while distrust refers to an individual who are self-closure of a relationship investment. According to Cho (2006), the undermining of trust does not represent distrust, and distrust is not synonymous with weak trust, since the antecedents of trust and distrust differ.

Although trust has been studied in various disciplines (Ferrin and Dirks, 2006) and is studied in different marketing contexts (Pavlou and Gefen, 2004; Luo, 2006), the empirical study of trust in computer mediated peer-to-peer environments remains limited (Vance *et al.*, 2008). Furthermore, most studies of trust have taken a cross-sectional approach (Mohd and Abdulla, 2006; Massey and Dawes, 2007; Gupta *et al.*, 2009) with relatively few time-series studies (Palmer and Huo, 2013).

As discussed above, trust is an incentive factor that motivates an individual to contribute in knowledge sharing (Chiu *et al.*, 2006) which facilitates the sustainability of an online community. It is proposed therefore in this thesis that the evolution of overall trust can encourage individuals to actively participate in online information sharing.

2. 3.5 The evolution of perceive critical mass within online forums

In this section, in order to explore the structure of online forums, studies in the field of network science are summarized. The dynamic aspects of network structural influence on the evolution of online forums are discussed. The role of the critical mass members in the evolution of online forums is finally explained.

2.3.5.1 Background

Online forums can be conceptualised as network graphs consisting of nodes (human actors) and edges (social relationships), and imply structures marked by the emergence of hubs (Dorogovtsev *et al.*, 2001; Albert and Barabási, 2002; Newman, 2003) which are influenced by mass voluntary collaborations (knowledge contributions) and interactions among members (message exchanges) (Wasco *et al.*, 2009; Westland, 2010).

The types of networks should influence the online connectivity, because the functionalities of networks vary with their type (Newman, 2005). Studying the differences in the connectivity among members (also the degree distribution) is the major method to distinguish the type of networks (Newman, 2005), with their degree distributions varying in Poisson, lognormal, stretched exponential and power-law (Clauset *et al.*, 2009).

Erdős and Rényi (1960) develop random network theory in which a member chooses to randomly connect to another with a given probability. As a result, the network contains on average $p \cdot N(N-1)/2$ edges. This is a typical binomial problem, members either connect or not to each other. That is, the probability which a member can have k connections (P_k) can be described as: $P_k = \binom{N-1}{k} p^k (1-p)^{N-1-k}$, the average connections is therefore, $\bar{K} = p(N-1)$. For a large N , i.e. the numbers of members or the size of network, the degree distribution follows the Poisson distribution: $P_k = e^{-\bar{K}} \bar{K}^k / k!$ (2.5), informing that the vast majority of members have roughly the same connections and a few members have connections which deviate significantly from the average (Barabasi and Frangos, 2014).

Randomised growth can also be described with the lognormal distribution. Lognormal distribution assumes random variables $Y_i = \ln X_i$, $i > 0$ have normal (Gaussian) distribution. The density function for a lognormal distribution satisfies (e.g. Clauset *et al.*, 2009): $f(x) = \frac{1}{x} \exp[-\frac{(\ln x - u)^2}{2\sigma^2}]$ (2.6), where u is the mean and σ is the standard variance (σ^2 is the variance). For example, the size growth or shrinkage of an organism depends on a random variable with its log10 converging to a normal distribution (Mitzenmacher, 2004).

Another distribution in the network analysis can be the stretched exponential distribution (also known as Weibull distribution), which has the form (e.g. Clauset *et al.*, 2009): $f(x) = x^{\beta-1} e^{-\lambda x^\beta}$ (2.7), where $\beta > 0$ and $\lambda > 0$ are the shape and scaling parameters accordingly. Stretched exponential distribution is often used to study the failure rate in a growth system; for example, email spam within a college (Newman 2005); when $\beta=1$, it is the exponential distribution and informs a constant failure rate; when $0 < \beta < 1$, it refers to a decrease failure rate because the defects occur earlier and are wiped out from the population; when $\beta > 1$, it is roughly the “S” shape curve (i.e. concave at Firstly, then convex after a point), suggesting an increase failure rate.

Although those distributions are highly peaked and skewed, they have finite variance around a mean value, i.e. the average connections (degree) associated to members. However, a non-negative variable that follows the power law distribution does not fit this pattern (Clauset *et al.*, 2009).

Members who are associated with a particular member within a network are called first-order of neighbours; neighbours of the first-order of neighbours are the second-order of neighbours (Barabasi, 2013). The power law distribution refers to variations between the numbers of the first and the second neighbours that tend to be infinite. There is a continuous hierarchic variation but no typical scale connectivity between members, and this is a so called scale-free network (Barabasi and Albert, 1999). The concept of scale-free suggests a power law tail, $f(x) = x^{-\gamma}$ (2.8), or power law with the exponential

cut off, $f(x) = x^{-\gamma} e^{-\lambda x}$ (2.9), characterising the distribution of membership connectivity marked by growth and preferential attachment (Newman, 2005). In (2.8) and (2.9), γ or λ represents the scaling parameter. For example, a web grows in time by attracting other links connecting to it, and the new jointed links prefer attaching to a web, depending on the previous level of links that the web has (Barabasi and Albert, 1999). Table 3 summarizes the formal expressions of the above discussed distributions.

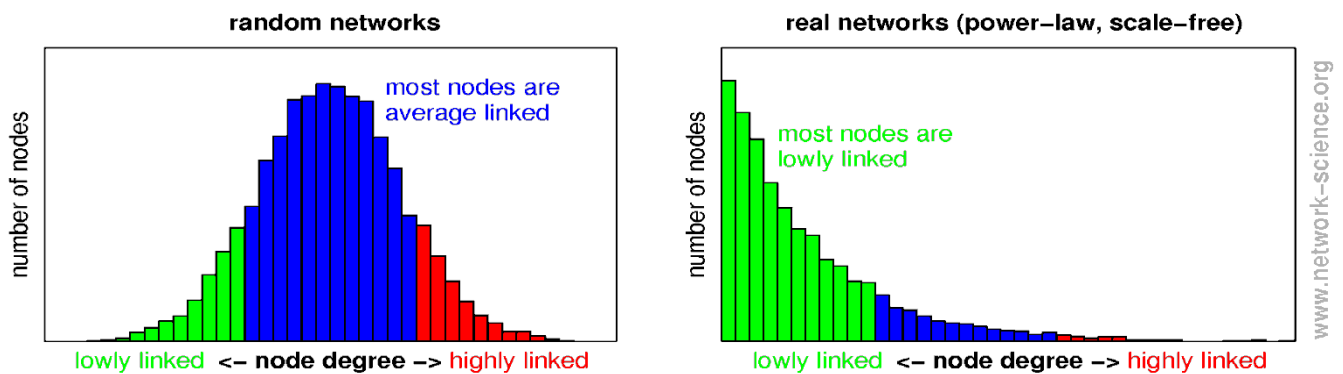
Table 3: Degree distributions applied to networks

Distributions (Continuous)	P(x) = Cf(x)	
	f(x)	C
Power law	$x^{-\gamma}$	$(\gamma - 1)x_{min}^{\gamma-1}$
Power law with exponential cut off	$x^{-\gamma} e^{-\lambda x}$	$\frac{\lambda^{1-\alpha}}{\Gamma(1-\alpha, \lambda x_{min})}$
Exponential	$e^{-\lambda x}$	$\beta \lambda e^{\lambda x_{min}^{\beta}}$
Stretched exponential	$x^{\beta-1} e^{-\lambda x^{\beta}}$	$\beta \lambda e^{\lambda x_{min}^{\beta}}$
Lognormal	$\frac{1}{x} \exp[-\frac{(\ln x - u)^2}{2\sigma^2}]$	$\sqrt{\frac{2}{\pi\sigma^2}} [\text{erfc}(\frac{\ln x_{min} - u}{\sqrt{2}\sigma})]^{-1}$

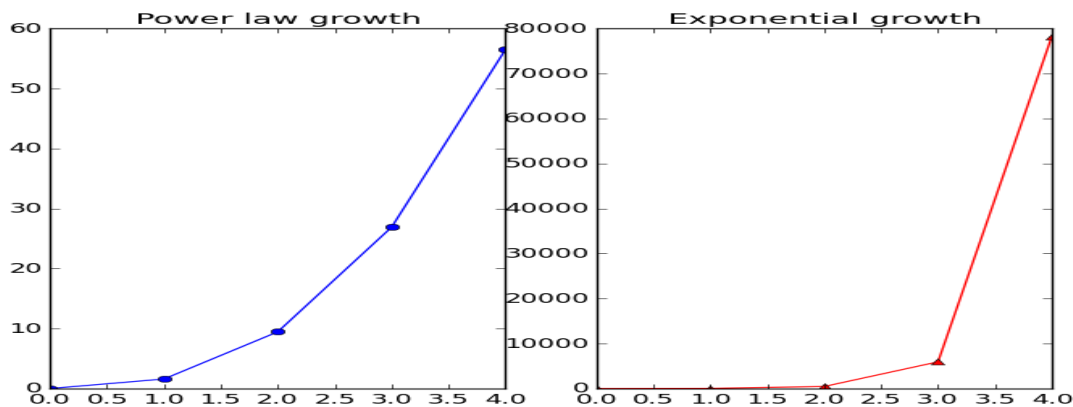
Clauset *et al.*, 2009

The concept of scale-free is developed by Barabási and Albert (1999) (“BA model”) from the random network theory (Erdős and Rényi, 1960) (“ER” model) and the small world network (Watts and Strogatz, 1988) (“WS” model). Both the Erdős-Rényi and Watts-Strogatz models share an assumption that there are fixed N members who are either randomly connected or reconnected with a probability p. As a result of which, their degree distributions are similar to the bell curves. While, the distance among members is shorter within a WS network than that with an ER network. However, as Barabasi and Albert (1999) point out: a real world network can enjoy a continuous increasing number of members throughout its lifetime. In many circumstances, members typically choose to connect with members rather than randomly. Stimulated by these observations, they developed the BA model, whose connectivity among members is characterized by the absence of the mean connections (such as a peak in the bell curve).

Figure 4: An example of the bell curve and the power law distribution and growth



Source: www.network-science.org



(Figure generated by Python)

Within scale-free networks, the majority of members have few connections but few members (hubs) hold a majority connection (Barabasi and Albert, 1999). Online forums provide an opportunity for members to exchange ideas with each other. Critical mass members defined in this study are those who contribute a majority of knowledge within online forums; they therefore should have more connections than the free-riders. In this sense, hubs and critical mass members have equivalent meanings in this study.

2.3.5.2 Barabasi and Albert's Model (BA model)

The concept of scale-free suggests the power law tail charactering the degree distribution (Barabasi and Albert, 1999; Newman, 2003; Dorogovtsev *et al.*, 2001; He *et al.*, 2009; Wang and Dai, 2009; Shi, 2011).

In order to illustrate the evolution of scale-free networks, Barabasi and Albert (1999) address the Barabasi-Albert (BA) model focusing on two generic attributes in many real world networks, exhibited through growth and preferential attachment. For example, the citations of an academic paper grow exponentially in time through new citation, and the new citation is more likely to depend on the cited times which may indicate the popularity of acceptance. These two properties demonstrate a good example of simplifying complex phenomena.

In the field of complex network science, nodes / vertexes are the points of intersections /connections within a network (e.g. Newman, 2005). Thus, members within online forums are called nodes or vertexes. The introduction to the logic of the BA model proposed by Barabasi and Albert (1999) is as follows:

Growth: starting with a small number m_0 of nodes, and a new node is introduced at every time step to m existed nodes, with $m \leq m_0$, in other words, this new node has m edges.

Preferential attachment: the probability Π that this new node will connect an existing $node_i$ depends on the degree k_i of $node_i$ and k_j of existed $node_j$, so that
$$\Pi_i = \frac{k_i}{\sum_j k_j} = \frac{k_i}{2mt} \quad (2.10).$$

Denoting $n(k,t)$, the numbers of nodes with degree k at time t , the degree distributions is defined:

$P_k(t) = \frac{n(k,t)}{N}$. After t time, the master equation for the expected nodes with degree k within indirect

network is expressed (Barabasi, 2013):
$$(n+1)p_k(t+1) = np_k(t) + \frac{k-1}{2} p_{k-1}(t) - \frac{k}{2} p_k(t) \quad (2.11),$$
 where the

numbers of nodes with degree k that obtain an additional link become nodes with degree $(k-1)$, $\frac{k}{2} p_k(t)$;

the numbers of nodes with degree $(k-1)$ that gain a new edge/link become nodes with degree k ,

$\frac{k-1}{2} p_{k-1}(t)$; the numbers of nodes with degree k that are not being linked with the new node is $np_k(t)$.

(2.11) calculates the degree distributions of all nodes with degree k ($>m$), and the computation represents

a recursive process. A mathematical relationship is obtained to describe the degree distribution in BA model (Bollobás *et al.*, 2001): $P(k) \sim k^{-\gamma}$ (2.12), where γ is the degree exponent that equals to 3.

\therefore

$$p_k = \frac{k-1}{2} p_{k-1}(t) - \frac{k}{2} p_k(t)$$

$$2p_k = -p_{k-1} - k[p_k - p_{k-1}] = -p_{k-1} - k \frac{p_k - p_{k-1}}{k - (k-1)} = -p_{k-1} - k \frac{\Delta p_k}{\Delta k}$$

(i.e. Starting from a stationary view:

\therefore

$$p_k = \frac{1}{2} \frac{\Delta[kp_k]}{\Delta k}$$

\Rightarrow

$$p_k \sim k^{-3}$$

BA model is a particular case in the region of scale-free networks (e.g. Clauset *et al.*, 2009). (2.12) describes the probability of nodes connecting to k others within scale-free networks, with γ that is typically setting between 2 and 3. As discussed in the section 2.7.1, the power law degree distribution suggests that a scale-free network evolves into the state where a few highly connected hubs can dominate the network's connectivity (Barabasi and Albert, 1999), and hubs have the equivalent meanings as the critical mass members. The following section 2.7.3 will further explain the role of critical mass members in sustaining online forums, and the method borrowed from the network science to identify the critical mass members.

2.3.5.3 Theory of critical mass applied in scale-free networks

The theory of critical mass proposed by Oliver and Marwell (1988) is often studied from the social perspective by exploring the influences of the perceived size of a network on individuals' behaviours (e.g. Cho, 2011; Shen *et al.*, 2013). Few studies (Westland, 2010; Centola, 2013) have sought to understand the theory of critical mass applied within networks. This is due to the difficulty of calculating the critical point over which a mass phenomenon emerges (Shen *et al.*, 2013). The critical point and the phase transition are two important aspects underlying the theory of critical mass (e.g. Oliver and Marwell, 1988). In this section, the method to address these two issues (borrowed from network science) is

discussed below. Firstly, it explains the concept of phase transition. Secondly, the method proposed by Cohen *et al.* (2002) to calculate the critical point is described. Finally, the role of the critical mass members in sustaining online forums is discussed.

2.3.5.4 Phase transitions

The discovery of scale-free networks where degree distributions obey the power law has evoked great interest in network research. In fact, the emergence of complex network theory is a topic about complex systems that has been studied in diverse fields from biology, physics to economics (Wang *et al.*, 2006). To date, there is no widely accepted definition of a complex system. However, the shared understanding in the field leads to the theory of self-organised criticality (SOC) (Bak *et al.*, 1987) which explains how the non-linear interactions among nodes can bring a phase transition even in the circumstance of non-central control, asymmetric information and local movement. The term phase transition often involves a system breaking (e.g. Wand and Dai, 2009). For example, phase transition is employed to explain the disappearance of the dinosaur (from existence to disappearance), avalanches (from small to immense in size), earthquakes (from linked to broken) and bull or bear stock markets (changes in opposite direction) (Wang and Dai, 2009). Each point in the former phase has the same proprieties, but differs from that in the phase changed after a critical point.

Self-organised criticality is considered as a ‘must’ connection to scale-free networks which is further supported by the observation of the power law tails (He *et al.*, 2008). The Sandal model by Bak *et al.* (1987) is the classical example to illustrate the concept of SOC. Consider a desk where a number of grains are introduced randomly. Thereafter a new grain is added and the height of sandal increases by repeating this step. As the pile grows, its slope becomes steeper until additional grain triggers a local avalanche. After a while, the pile reaches a critical height, newly joined grain leads to a large avalanche and the pile spreads aside. A critical point occurs where a single movement, such as falling of a new grain at this point can evoke the change of the whole system. This is self-organised in that there is no invisible

hand or external factors that are manipulating this phase transfer, because this pattern occurs spontaneously.

Results of extensive simulations on the sandal model (Albert and Barabasi, 2000) indicate that the distribution of the size of the avalanche follows the power law form and is described as : $N(s) = s^{-t}$, where $N(s)$ stands for the size of the collapse event, t is the exponent parameter. The negative of t indicates that the size of collapse decreases when the size of sandal S increases. The idea is that local disturbances in time and space are accumulated exponentially, which allows the occurrence of the giant component at the critical point over which the avalanche is generated.

The concept of phase transition in the study is explained. A phase transition is associated with the percolation theory that seeks to understand the cluster structure within a network (Erdős and Rényi, 1960; Burton and Keane, 1989; Molly and Reed, 1995; Cohen *et al.*, 2003; Newman, 2005). Percolation theory is often studied in an infinite system corresponding to the emergence of the mass phenomenon after the critical point. As a general rule, the presence of a giant component in a random network or spinning cluster in a scale-free network can ensure the ongoing growth of that network (e.g. Cohen *et al.*, 2002), which is further discussed below.

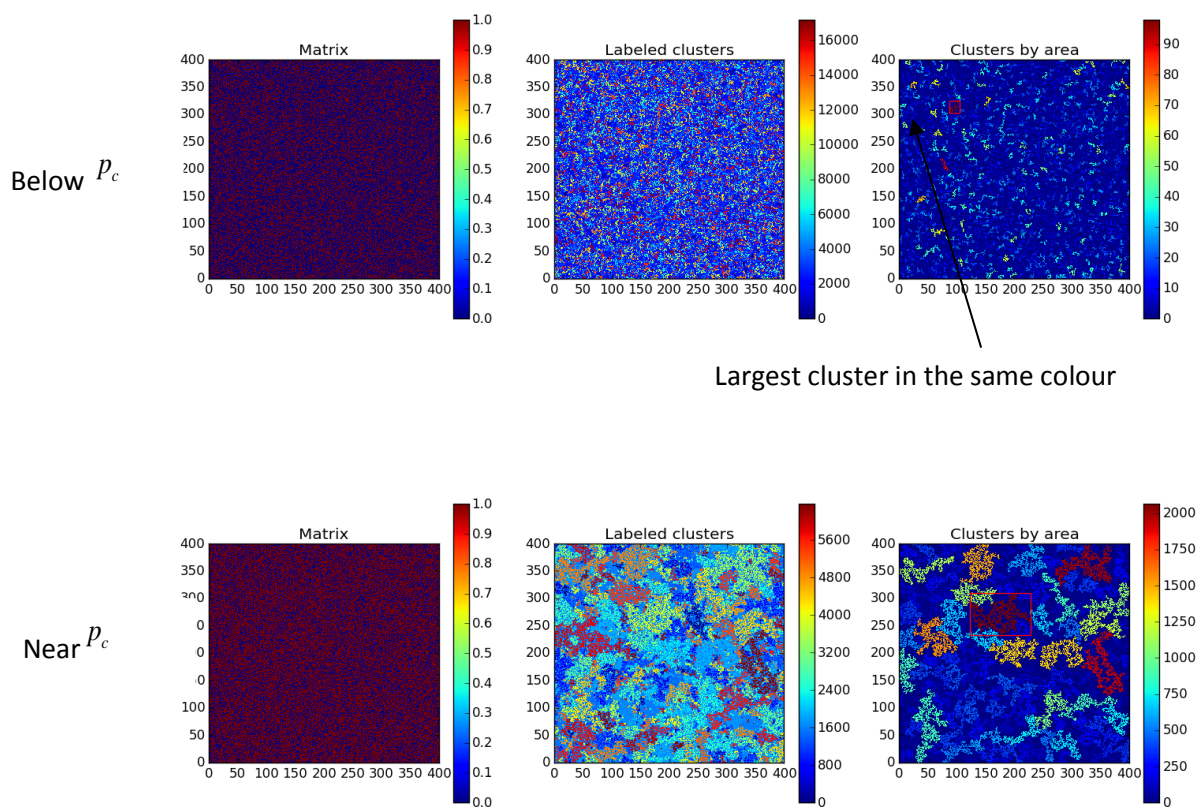
2.3.5.5 Percolation theory

Percolation theory studies an ongoing growth system by examining the cluster structure within it (e.g. Cohen *et al.*, 2002; Newman, 2005). As one arbitrarily chosen node belongs to the spinning cluster, for example, it can connect to everyone else in the network (e.g. Newman, 2005). In other words, there is always the path in the network that connects everyone through the spinning cluster (e.g. Barabasi and Frangos, 2014), or there is always new members joining to scale-free network, and potentially connecting to others through the spinning cluster.

To illustrate the concept of spinning cluster, a simulation on square (400 x 400 for a relative large size of network) is undertaken, with codes cited by 计算机博士 from Baidu Baike and Github (see figure 5 and

6). Far below the critical point p_c , there are many disconnected clusters whose size are not influenced by the network size. The largest cluster size rapidly increases approaching to p_c and the shape of clusters is roughly the same. Spinning cluster which is only limited by the network size emerges after the critical point, followed by an increase of the cluster area and a decrease of the numbers of cluster. The tail (after the critical point) of the numbers of cluster size follows the power law. With the emergence of the spinning cluster, the network transforms the phase from disconnect to connect. This is similar to the phase transition in the theory of critical mass (e.g. Roger, 1995) that the size of new technology adopters expands over a critical point.

Figure 5: Largest cluster and numbers of clusters



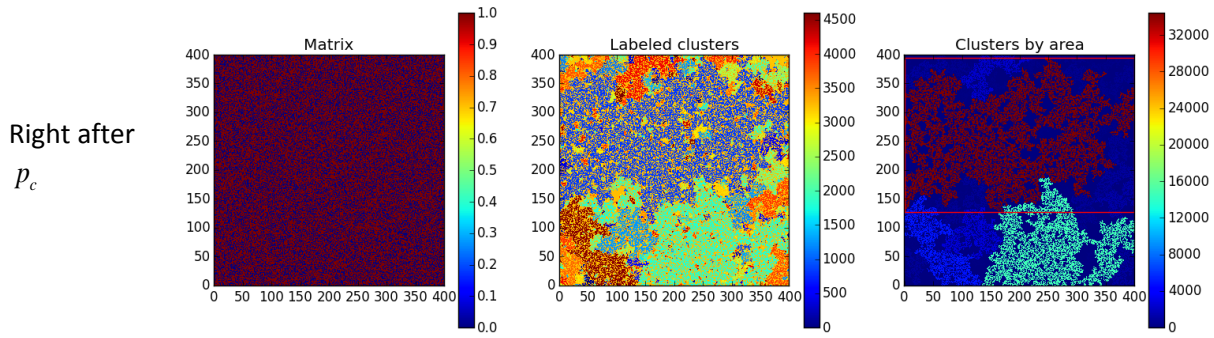
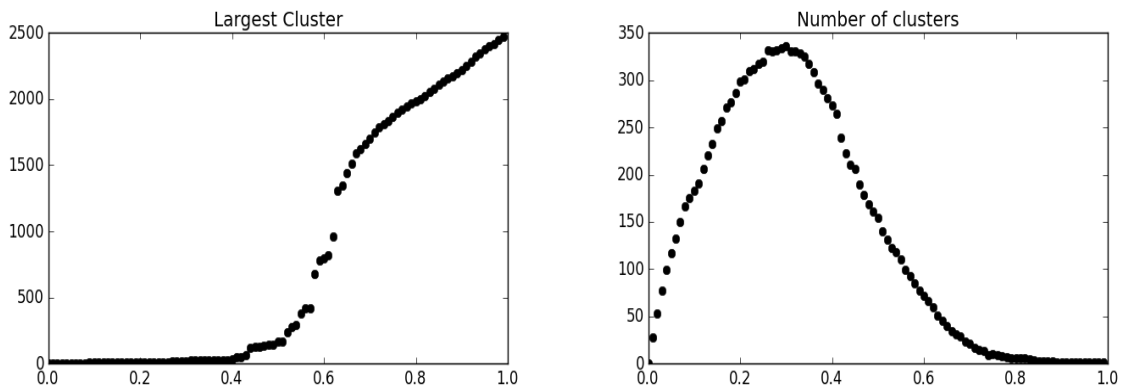


Figure 6: Largest cluster and number of clusters



For a square lattice with 2 dimensions, the critical point is known at 0.5927... (Newman, 2005).

The percolation phase transitions are geometric which involves the parameters such as network size and heterogeneous systems (e.g. Newman, 2005), similar to the important assumption of heterogeneous population for the theory of critical mass (Oliver and Marwell, 1985). Defining $p(s)$ is the probability of a randomly chosen node that belongs to the largest cluster with size s . The distribution of $p(s)$ is therefore dimensionless but s is an area with some unit measure, denoting a (Newman, 2005). The dimensionless

distribution function $p(s) = C f\left(\frac{s}{a}, \frac{a}{\langle s \rangle}\right)$ (2.13), where C is a normalized constant to ensure $\sum_s p(s) = 1$.

This is because after the critical point, the mean area of the largest cluster size $\langle s \rangle \rightarrow \infty$ by the definition

of spinning cluster, $p(s) = C f\left(\frac{s}{a/b}\right) = \left(\frac{C'}{C}\right) p(bs) = g(b)p(s)$ (2.14), where b is something to rescale a so

that the cluster size and shape are roughly the same (coarse graining technique), and $\frac{C'}{C}$ represents $g(b)$.

Power law is the only solution satisfying $p(bx) = g(b)p(x)$ (2.15) (Newman, 2005). Thus, a scale-free network is self-sustaining (e.g. Barabasi, 2013).

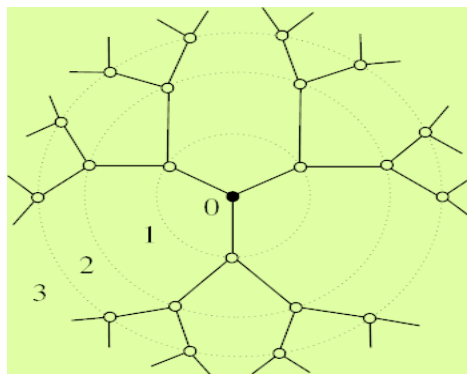
2.3.5.7 Critical point

Previous numerical analytical results (Albert and Barabasi, 2000) derived from percolation processes in a Cayley tree indicates that the condition for the emerging of the giant component (in random network) is:

$p_c = \frac{1}{z-1}$ (2.16), where p_c is the critical point when the giant component is emerging. z ($=3$) is the maximum acquaintance of the origin site. $z-1$ means that at least one of $z-1$ edges connects the origin node to others. That is, a Cayley tree with coordination number equals to three, whose average degree approaches to two. A Cayley tree holds a property in common with a random graph (Albert and Barabasi, 2002).

Similar to the theory of self-organised criticality that describes the phase transition happening almost immediately after the occurrence of the critical components, linkage possibility $p \geq p_c$ assures that network growth is continuous or self-sustaining since the giant component appears and they own infinite outgoing edges.

Figure 7: An illustration of the Cayley tree



Source: Westland (2010)

With regard to the scale-free network, it always percolates, and there is no critical point after which the spinning cluster emerges (Barabasi, 2000; Dorogovtsev and Mendes, 2001; Cohen *et al.*, 2002). However, for the finite scale-free network, no matter how big in size it can be, it is finite with empirical data (e.g. Barabasi and Frangos, 2014), and there exists the critical point (e.g. Cohen *et al.*, 2002).

Similar to an epidemic spreading within networks (e.g. Pastor-Satorras and Vespignani, 2001); members within online forums can be classified into three categories: members who are removed so that it is not possible for them to invite new members to join in; members who are susceptible to leaving online forums; and members who steadily contribute within online forums. Defining ν is the probability that members are moving from the state of susceptible to removed; δ is the probability that members are moving from the state of removed to susceptible; λ is the spreading critical point and expressed as (e.g.

Pastor-Satorras and Vespignani, 2001a) : $\lambda = \frac{\nu}{\delta}$ (2.17).

Defining the density probability that nodes with degree k at time t are susceptible, $\rho_k(t)$, the following equation is given embedded in the proposition by Pastor-Satorras and Vespignani (2001) :

$$\frac{\partial \rho_k(t)}{\partial t} = -\rho_k(t) + \lambda k [1 - \rho_k(t)] \Theta(\rho_k(t)) \quad (2.18) \quad \text{where } \Theta(\rho_k(t)) \text{ is the probability of edges to the nodes}$$

with k connections at time t who are susceptible. The right side of (2.18) equals zero, thus,

$$\rho_k = \frac{k \lambda \Theta(\lambda)}{1 + k \lambda \Theta(\lambda)} \quad (2.19) . \text{ This is because that } \rho_k(t) \ll 1 \text{ which can be omitted, and } \lim_{t \rightarrow \infty} \rho_k(t) = \rho_k . \text{ For the}$$

scale-free networks, the probability of edge to nodes with degree s is (e.g. Barabasi and Frangos, 2014):

$$\frac{sP(s)}{\langle k \rangle} \quad (2.20) . \text{ In (2.19), } \Theta(\lambda) = \frac{\sum_k k P(k) \rho_k}{\langle k \rangle} \quad (e.g. \text{ Wang } et al., 2006) \quad (2.21) . \text{ Combining (2.19) and}$$

$$(2.21), \text{ it obtains (e.g. Wang } et al., 2006) : \quad \Theta = \frac{1}{\langle k \rangle} \sum_k k P(k) \frac{\lambda k \Theta}{1 + \lambda k \Theta} \quad (2.22) . \text{ For } \Theta \neq 0 , \text{ it requires}$$

$$(e.g. \text{ Pastor-Satorras and Vespignani, 2001}) : \quad \frac{d}{d\Theta} \left(\frac{1}{\langle k \rangle} \sum_k k P(k) \frac{\lambda k \Theta}{1 + \lambda k \Theta} \right) \Big|_{\Theta=0} \geq 1 \quad \Rightarrow$$

$$\frac{1}{\langle k \rangle} \sum_k k P(k) \lambda k = \frac{\langle k^2 \rangle}{\langle k \rangle} \lambda \geq 1 \quad (2.23)$$

From (2.23), the critical point over which a spinning cluster emerges $\lambda_c = \frac{\langle k \rangle}{\langle k^2 \rangle} = \frac{\langle k \rangle}{\langle k(k-1) \rangle}$, subject to $\lambda \succ \lambda_c$ (2.24). With (2.24), a scale-free network whose scaling parameter setting between 2 and 3 ($2 < \gamma \leq 3$), $\lambda_c \rightarrow 0$, subject to $N \rightarrow \infty$ and $\langle k^2 \rangle \rightarrow \infty$ (e.g. Wang *et al.*, 2006).

2.3.5.8 The importance of critical mass members in sustaining online forums

For scale-free networks (see the section 2.7.1), hubs who own the majority connections of networks are widely recognized as the key factor of interest (He *et al.*, 2009). This aspect is crucial against random attack to the scale-free networks (e.g. Barabasi, 2000). That is, if hubs are not attacked, the network remains largely undamaged. In contrast, a scale-free network is vulnerable to the attending attack of hubs (Albert *et al.*, 1998; Jeong *et al.*, 2000; Newman, 2003). Random attacking scale-free networks refer to randomly removing members in order to destroy the spinning cluster, so that the self-sustaining property is no longer ensured. Attending attack refers to removing hubs or members / nodes in an increasing fraction in order to break down the network (Albert *et al.*, 2000). As discussed above, hubs or critical mass members reflect the connectivity that a real network allows, and they share the similar meaning. That is, the role of critical mass members in sustaining online forums is able to be reflected through the robust and fragility characteristics of scale-free networks.

Defining f is the fraction of randomly attacked nodes against the total numbers of N nodes within a scale-free network, the probability for a node with degree k_0 becoming k ($k \leq k_0$) is (Cohen *et al.*, 2000):

$$C(k_0, k)(1-f)^k f^{k_0-k} \quad (2.25)$$

Therefore, the degree distribution of the network after random attacking

$$P_{new}(k) = \sum_{k_0=k}^{\infty} P(k_0) C(k_0, k) (1-f)^k f^{k_0-k} = \sum_{k_0=k}^{\infty} P(k_0) \left(\frac{k_0!}{k!(k_0-k)!} \right) (1-f)^k f^{k_0-k} \quad (2.26), \quad \text{where}$$

$\langle k \rangle_{new} = \langle k_0 \rangle (1-f)$ (2.27), and $\langle k^2 \rangle_{new} = \langle k_0^2 \rangle (1-f)^2 + \langle k_0 \rangle f(1-f)$ (2.28) (e.g. Cohen *et al.*, 2003).

(For example,

$$\langle k \rangle_{new} = \sum_{k=1}^{\infty} k \sum_{k_0=k}^{\infty} P(k_0) C(k_0, k) (1-f)^k f^{k_0-k} = \sum_{k_0=1}^{\infty} P(k_0) \sum_{k=1}^{k_0} k C(k_0, k) (1-f)^k f^{k_0-k} = \sum_{k_0=1}^{\infty} k_0 P(k_0) (1-f) = \langle k \rangle (1-f) ;$$

$$\langle k^2 \rangle_{new} = \langle k_0 \rangle (1-f) \sum_{k_0=1}^{\infty} P(k_0) \sum_{k=1}^{k_0} k C(k_0, k-1) (1-f)^k f^{k_0-k+1} = [\langle k \rangle (1-f)]^2 + \langle k \rangle (1-f) f$$

The critical fraction of spinning cluster within scale-free network is (Cohen *et al.*, 2002):

$$f_c = 1 - \frac{1}{\frac{\langle k^2 \rangle}{\langle k \rangle} - 1} \quad (2.29), \text{ where } \frac{\langle k^2 \rangle}{\langle k \rangle} \text{ measures the divergence between the average of the second order}$$

of neighbours (second movement, neighbours of neighbours) and the first order of neighbours (first movement, the nearest neighbours (Barabasi and Albert, 1999). By incorporating the (2.26) into (2.27),

$$\text{subjective to the condition (e.g. Dorogovtsev et al., 2001; Moreira et al., 2002): } k_{\max} \sim k_{\min} N^{\frac{1}{\gamma-1}} \quad (2.30),$$

the critical fraction of nodes to be randomly removed within scale-free networks becomes:

$$f_c^{random} = 1 - \frac{1}{\frac{\gamma-2}{\gamma-3} k_{\min}^{\gamma-2} k_{\max}^{3-\gamma} - 1} \quad \text{or} \quad k = \frac{\langle k^2 \rangle}{\langle k \rangle} = \frac{(2-\gamma)}{(3-\gamma)} \frac{(k_m^{3-\gamma} - m^{3-\gamma})}{(k_m^{2-\gamma} - m^{2-\gamma})} \text{ for } 2 < \gamma \leq 3 \quad (2.31). \text{ When } k_{\max} \rightarrow \infty$$

in $N \rightarrow \infty$, thus $f_c \rightarrow 1$. In other words, this needs to randomly remove almost all nodes within scale-free networks to fragment networks (Barabasi, 2013). This property demonstrated within scale-free networks, is what Barabasi (1999) calls, robustness in random failure. (2.30) is obtained with empirical studies (e.g. Dorogovtsev and Samukhin, 2001; Moreira *et al.*, 2002).

On the other hand, scale-free network is fragile when network attacks start by removing the hub with the highest degree (Barabasi, 1999). Similar to random attack, there are two major consequences of attending

attack, i.e., the maximum degree k_{\max} decays to k'_{\max} ($k'_{\max} < k_{\max}$), and the degree distribution $P(k)$

changes to $P'(k)$, and the network is no longer scale-free because hubs hold a majority of connections of a network (e.g. Cohen *et al.*, 2001). The critical fraction in the case of intentional attack is given (Cohen *et al.*, 2001): $f_c^{attending} = f^{\frac{2-\gamma}{1-\gamma}}$ (2.32) , and $k'_{max} = k_{min} f^{\frac{1}{1-\gamma}}$ (2.33) . For $\gamma=2$, $f_c^{attending} \rightarrow 1$, with a very small critical fraction, the network is becoming fragmented into many disconnected components (e.g. Cohen *et al.*, 2001). This is the fragility of intentional attacks of scale-free networks (e.g. Barabasi, 2013).

This section has provided a background about how the theory of critical mass can be applied online. Previous studies have mainly investigated the phenomenon after the critical point (e.g. Shen *et al.*, 2013), leaving the dynamic process regarding how the mass phenomenon emerges little researched. Study three addresses this limitation, in particular, by examining the critical point role that critical mass members play in the evolution of online forums seen as networks.

2.4 Summary of chapter

Online forums have been recognized as an important and efficient tool for firms who seek to communicate with customers (Vargo and Lusch, 2008). However, the issue of “free-riders” makes sustaining online forums very challenging (Wasco *et al.*, 2009). Previous research in sustaining online forums is rare (Ridings and Wasco, 2010), with few exceptions that have different focuses on online knowledge continuance. Those findings include interpersonal trust (He and Wei, 2009), institutional trust (Zimmer *et al.*, 2010) and the theory of critical mass that can explain the evolution of knowledge contributions within online forums (Wasco *et al.*, 2009).

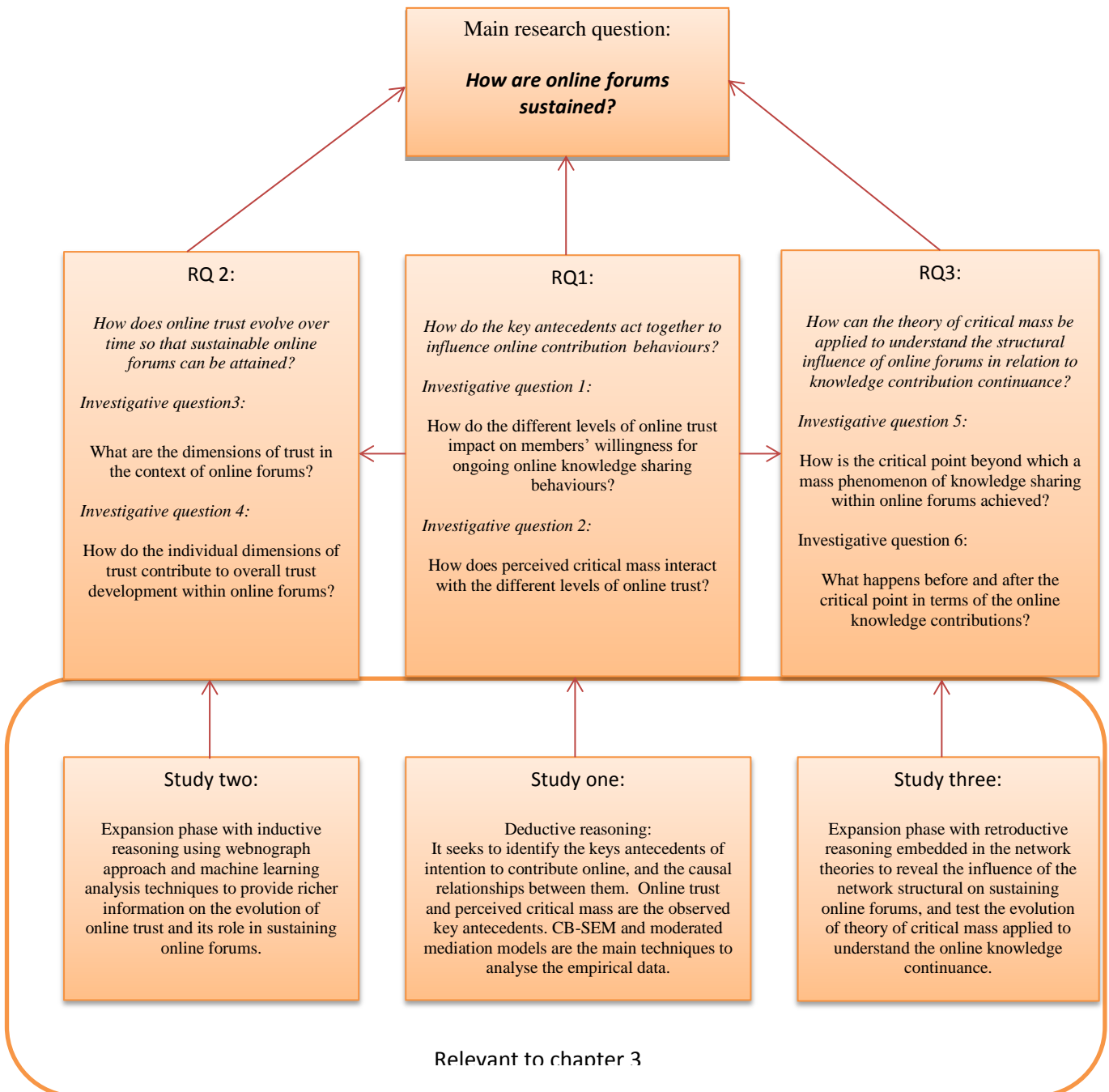
Yet, existing studies have not sought to investigate the dynamic aspects of the identified antecedents; How these antecedents act together and play a role in sustaining online forums remains under studied. The principle contribution of study one has been to investigate how the different antecedents that are dynamic in nature impacts together on the determinants of online voluntary contributions, and the causal relationships between the identified antecedents. However, the dynamic aspects of trust and critical mass concepts are explained descriptively in study one. Study two seeks to add knowledge in the field of trust

studies by investigating how it is evaluated and undermined in the context of online forums. Critical mass theory involves a phase transition after which a mass phenomenon occurs (Marwell *et al.*, 1988). However, previous studies of critical mass have mostly focused on what happened after the critical point (Cho, 2011; Shen *et al.*, 2013), with few studies (Westland, 2010; Centola, 2013) having defined the formal matrix of critical mass. Study three seeks to fulfil the knowledge gap by examining the critical mass process within online forums. Both study two and study three are extension research to study one. This thesis takes a holistic view in order to provide a complete understanding of ongoing knowledge contribution behaviours in the context of online forums.

Chapter 3 Methodology

Before a further discussion of the methodology used in this thesis, the rationale of the thesis explained in the section of “Introduction” is recalled in Figure 3.1:

Figure 3.1: Research design: rationale of this thesis



3.1 Epistemological overview

The methodological approach used in this study embraces mixed methods which integrates quantitative and qualitative studies. According to Tellis *et al.* (1999), results generated from only one type of research method is often biased because only issues amenable to that method are considered. The use of more than one research method is suggested in order to provide a more robust insight into a complex research question because it can triangulate findings across methods (Jick, 1979).

Mixed-methods research designs that combine the findings generated from different research have become increasingly popular (Greene 2007), and recognised as a superior research method to develop understanding around a phenomenon (Jick 1979). Despite the claimed strengths of mixed methods research, it has been noted that this approach isn't often employed in marketing research (Golicic and Davis, 2012) and information system (IS) studies (Venkatesh *et al.*, 2013).

Another value of mixed methods research is that it allows the researcher to address both confirmatory and exploratory research questions simultaneously (Venkatesh *et al.*, 2013). In the field of IS and social science research, qualitative research has typically been conducted for exploratory research in order to develop deep understanding of emerging concepts or phenomena (Punch 1998). In contrast, quantitative research has been applied for a confirmatory research approach, typically in order to test causal relationships between constructs within a proposed structural model and to test hypotheses (Venkatesh *et al.*, 2013). However, a criticism of employing a single method is that researchers may not be able to develop or extend substantive theory in a rich way (Mingers, 2001; Venkatesh *et al.*, 2013). Venkatesh *et al.* (2013) emphasize that mixed methods research is a substantial/or integrative method that can discover components and unveil interrelationships among components related to a phenomenon.

Mixed methods research which uses multiple methods (Venkatesh *et al.*, 2013) has been termed the third methodological movement (paradigm) whereby quantitative research methods represent the first movement, and qualitative research methods represent the second movement (Zhang, 2011; Venkatesh *et al.*, 2013). According to Venkatesh *et al.*, (2013), the terms "multiple methods" and "mixed- methods"

researches that have been used interchangeably by many researches should be considered distinct. In the multiple methods approach, a researcher employs more than one method but is restricted to a single worldview; for instance, diverse research methods within an exclusively qualitative or quantitative paradigm. In contrast, mixed- methods research involves multiple worldviews, e.g. in line with methodological combination of qualitative and quantitative paradigms. In other words, all mixed-methods research encompasses multiple research method, while all multiple research approaches does not constitute not mixed- methods research (Venkatesh *et al.*, 2013).

Mixed- methods research is of particular value when the objective of a research task is to provide a holistic view of phenomena that has only been partially understood (Venkatesh *et al.*, 2013). Within the context of online knowledge contribution, previous research has been conducted to understand antecedents related to either social or structure dynamics. Yet, there have been no significant published studies that have developed multifaceted insights in this context that have simultaneously used a combination of different types of knowledge which exist in the domains of sociology and physics. This thesis will adapt the mixed - methods research approach, involving both quantitative and qualitative methods to investigate voluntary contribution behaviours that have an effect on the sustainability of online forums (Wasco *et al.*, 2009).

This thesis is conducted embedded in a mixed- methods research that reflects an integration of conclusions derived from different studies. Structural equation modelling (SEM) is the analysis technique used to test data collected from survey that is designed for examining hypotheses defined in chapter two. In order to give an insight to the findings from the survey, one qualitative case study is conducted with the purpose of exploring the evolution of online trust that cannot be addressed by a survey. Another study, following retroductive approach, deals with the structural mechanism in explaining the theory of critical mass applied in the context of internet websites. Both study two and study three that take a dynamic view on antecedents of intention to online contributions, can provide further insights to the results generated by the deductive testing of the hypotheses. The following sections deal firstly with a

summary of the general characteristics of mixed- methods research. Subsequently the research design for each study is explained.

3.2 Inductive, deductive and retroductive research approaches

In terms of the possible methodological processes that can be adopted for this thesis, Blaikie (2000; cited in Baker, 2003) identify four main research methodologies: namely inductive, deductive, retroductive, and abductive approaches.

A definition of the *inductive* research approach provided by Blaikie (2000, p.102; cited in Baker, 2003), states that this methodology implies that "...meticulous and objective observation and measurement, and the careful and accurate analysis of data, are required to produce scientific discoveries." The *deductive* research approach on the other hand, is defined by Baker (2003, p. 124) as involving: "...the statement of a hypothesis and the conclusion drawn from it, the collection of appropriate data to test the conclusion and the rejection or corroboration of the conclusion." While *retroductive* research strategy also begins with an observed regularity, it attempts to discover a different type of explanation, with the explanation in this case being defined by Blaikie (2000, p. 25) as being achieved by: "...locating the real underlying structure or mechanism that is responsible for producing the observed regularity...Retroduction uses creative imagination and analogy to work back from data to an explanation."

The other methodological approach identified by Blaikie (2000) is not deemed suitable for this thesis. The abductive approach is more associated with a range of interpretivist experimental approaches that are better suited to the social sciences. Hence, the idea behind the methodological approach, according to Blaikie (2000, p. 114; cited in Baker, 2003) describes this approach as: "...the process used to generate social scientific accounts from social actors' accounts for deriving technical concepts and theories from lay concepts and interpretations of social life. "

The deductive research approach is served for study one that has sought to test a set of hypotheses embedded in the known theory, i.e. DTPB, and to abstract conclusions against observations (Snieder and

Larner, 2009). The deductive approach is often seen as “reasoning from general to particular” (Pelissier, 2008, p. 3). In other words, the conclusions from study one can provide an overall picture of the key antecedents affecting online intentional knowledge contribution behaviours.

The inductive research approach can be related to grounded theory. Baker (2003, p.160) refers to grounded theory when discussing ethnographic studies, and states that:

“When discussing the basic difference between a positivistic or interpretivistic (Phenomenological) approach to research, grounded theory was identified as an example of the latter with theory evolving from observation of a phenomenon. Such theorising might be limited to a specific relationship – a substantive theory – or be generalised to embrace a class of relationships through the statement of a general theory.”

Thus a methodology based in webnography is employed for study two, and this is considered particularly appropriate as this is a form of research that closely involves consumers who are highly involved in contributions to online forums. In grounded theory, information is recorded as it emerges from the study, with the aim then being for the theory to emerge from a systematic analysis of the data collected from the observations. As epitomised by Baker (2003, p. 160): “... grounded theory seeks to derive structure through the analysis of non-standardised data, while surveys define a structure and collect standardised data to enable the testing of hypotheses on which the structure is founded.”

Therefore, aspects of grounded theory also form the basis of the research strategy for study two of the thesis. However, the grounded theory used is not highly-structured or systematic in nature, due to the time constraints encountered in conducting the research. It is therefore found to be more suited to the explanation phase of the thesis. Unlike grounded theory in its purest form (Stauss and Corbin, 1998) in which the researcher begins the study with no pre-conceived ideas about the object or person under examination (Baker, 2003), it does not start with a completely blank sheet, but incorporates previously discovered knowledge.

The retroductive phase is related to the visualisation of an online forum as a network that is embedded in graph theory in study three. The mechanisms underlying network structures that focus on the role of critical mass members in sustaining discussions within online forums are in agreement with one of the results suggested by study one, which highlights the important role of a small group of contributors in evoking a mass contribution. This is observed and explained. Graph theory is a study of graphs that is “a way of specifying relationships among a collection of items” (Easley and Kleinberg, 2010, p. 37). Furthermore, Easley and Kleinberg (2010, p. 38) state that “graphs are useful because they serve as mathematic models of network structures,” and are applied in social network analysis and physics.

The use of different studies in this thesis, as opposed to multiple forms of just qualitative or quantitative research methodologies is deemed to be critical in establishing a true mixed-methods approach, which allows for a cross comparison of the different types of data, and provides a means of validating the findings of each study. The findings from the three types of data collection could also be compared with the literature review in order to more fully analyse and validate the results of each of the three studies.

3.3 Research design

As discussed in the previous section, mixed methods can provide a stronger approach for investigating an emerging phenomenon than can be provided by a single worldview (Teddlie and Tashakkori 2003, 2009). However, it does not lead automatically to the discovery, extension or development of a substantive theory (Venkatesh *et al.*, 2013). According to Mingers (2001), overcoming considerable barriers such as differences between culture and the physical location is an important challenge to researchers who intend to conduct mixed-methods research. Venkatesh *et al.* (2013) explain that mixed-methods research approaches should be sensitive to different research purposes. An awareness of such difference of purposes can help researchers to better understand the goals and outcomes of research. Table 4 summarizes the diversity of purposes that should inform the nature of mixed methods research.

Table 4: Purposes of mixed-methods research

Purposes	Description	Illustration
Complementarity	Mixed methods are used in order to gain complementary views about the same phenomena or relationships.	A qualitative study is used to gain additional insights on the findings from a quantitative study.
Completeness	Mixed methods designs are used to make sure a complete picture of a phenomenon is obtained.	The qualitative data and results provide rich explanations of the findings from the quantitative data and analysis.
Developmental	Questions for one strand emerge from the inferences of a previous one (sequential mixed methods), or one strand provides hypotheses to be tested in the next one.	A qualitative study is used to develop constructs and hypotheses and a quantitative study is conducted to test the hypotheses.
Expansion	Mixed methods are used in order to explain or expand upon the understanding obtained in a previous strand of study.	The findings from one study (e.g., quantitative) are expanded or elaborated by examining the findings from a different study (e.g., qualitative).
Corroboration/ Confirmation	Mixed methods are used in order to assess the credibility of inferences obtained from one approach.	A qualitative study is conducted to confirm the findings from a quantitative study.
Compensation	Mixed methods enable compensation for the weaknesses of one approach by using the other.	The qualitative analysis compensates for the small sample size in the quantitative study.
Diversity	Mixed methods are used with the hope of obtaining divergent views of the same phenomenon.	Qualitative and quantitative studies are conducted to compare perceptions of a phenomenon of interest by two different types of participants.

Adapted from Greene *et al.* (1989), Tashakkori and Teddlie (2003a, 2008), and Venkatesh *et al.*, 2013.

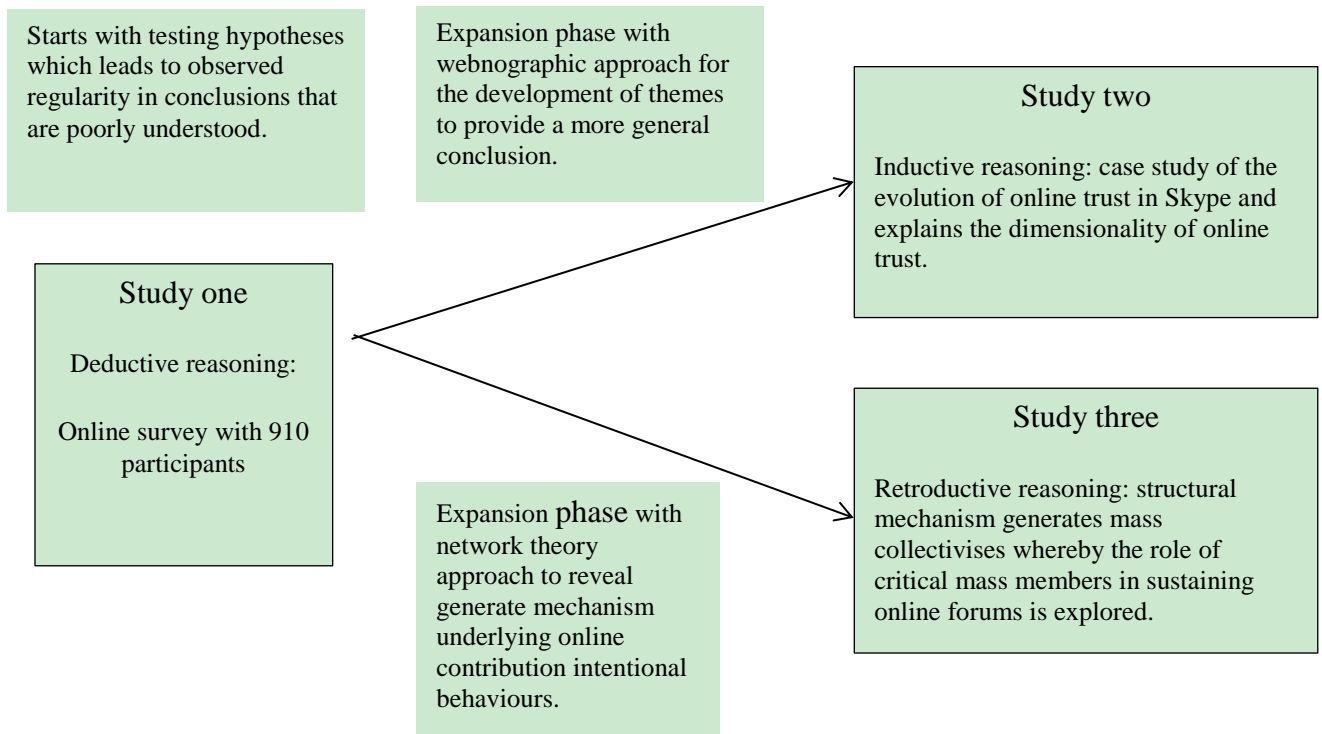
As discussed above, the main research aim of this thesis is to explore antecedents of sustainability in online forums, with three main investigative questions to explore antecedent influencing factors related to social and structure dynamics. An integrated model has been proposed following the outcome of the literature review, with hypotheses to test the relationships between constructs, notably trust, perceived critical mass and intention of knowledge sharing. However, the phenomena of development of trust and the structural mechanism that generates the mass collectivises are little understood in previous research, which should be further explored. Embedded in the discussion above, it is proposed that the findings from the quantitative research will be explored with substantial studies, i.e. study two and three.

According to Churchill and Iacobucci (2005), quantitative research can be either descriptive (establishing the frequency and basic relationship between variables), or causal (determining the existence of “cause-and-effect” relationships). This thesis will adopt firstly a descriptive research approach in order to establish the characteristics of online voluntary contribution intention in terms of attitudes, perceived behaviour control and subjective norms. Embedded in the results generated from descriptive quantitative study, mediation and moderated effects are investigated to understand the causal relationships between antecedents of intentional online contribution. Surveys are often employed in the descriptive research as they can provide an overview of key variables of interest to the researcher (Churchill and Iacobucci, 2005). In the case of online forums, an online survey was sent to internet users who join the activities of online forums. It is noted that the online survey used in this thesis is cross-sectional, and the data gathering from a sample of population is during a three week period (Churchill and Iacobucci (2005).

An online survey was developed to test the hypotheses. The quantitative stage is followed by two further substantial studies, which are conducted to provide a rich explanation about how trust evolves online and how critical mass members play an important role in the expansion of online forums. Chronologically, the research process is summarized in figure number 2.

An online survey is conducted in study one. Study one provides a deductive approach that explores social and structure dynamics in sustaining online forums. Following this, the data from study one are analysed to inform studies two and three. The content analysis of five years of online reviews for a telecommunications service provider is employed in study two, which is designed to understand the development of online trust. In study three, network analysis techniques are used to seek explanation for the influence of the structural dynamic of online forums on critical mass members. Findings from study two and study three seek to provide further explanation of results found in study one. Findings of each of the three studies are reported sequentially but are also embedded in the reporting of other stages of the study as the data emerges. The following sections will provide more details of methodology in each stage of the study.

Figure 8: Chronology of the research processes: mix data analyses integrated in mixed-methods



3.4 Methodology of study one – online survey

This section will examine the methodology of the online survey used in study one that is designed to test the framework proposed in the literature review. It starts by discussing construct development followed by an explanation of the sampling process, data collection method and the data analysis technique.

3.4.1 Online survey design

3.4.1.1 Identification of constructs

According to Hair *et al.* (2010), constructs should be identified before developing measurement scales.

The framework proposed in Chapter two has presented seven constructs with hypotheses relating to potential relationships between them. Table 5 summarizes the constructs identified in this study.

Dependant variables can be understood as the consequence of the independent variables (causes). Although the variables “attitude”, “perceived behavioural control” and “subjective norms” can be both the consequences and causes in the proposed model, they are variables with a mediating effect between “trust in members”, “trust in online communities”, “perceived critical mass” and “intention to sharing knowledge online”. Therefore, they are independent variables that can cause the dependant variable.

Table 5: constructs within the online survey

Dependent variable	Intention to share knowledge online
Independent variables	Attitude, Subjective norms, Perceived behaviour control Trust in members Trust in online forums Perceived critical mass
Demographic / control variables	Age Education Gender

3.4.1.2 Measurement scale development

Measurements can be understood to be either formative or reflective measures (Diamantopoulos and Winklhofer, 2001). Formative measures describe an observed variable that has a causal influence on a latent variable; while reflective measures assume that the latent variable causes the observed variables. For this research, formative measures are used because each construct is measured through multiple indicators.

Hair *et al.* (2010) explain that the response form for each multi-item scale is important to perform statistical data analysis and SEM efficiently. Likert scales can allow respondents to express directly the direction and the strength of their opinion (Garland, 1991). In fact, Likert scales (known also as rating scales) are widely employed in marketing research (Dawes, 2008). One important consideration when incorporating Likert scales into a study is to decide the number of points in a scale (Hair *et al.*, 2010). Guy and Norvell (1977) propose that scales should have a mid-point. This method ensures that

participants are not forced to choose either positive or negative opinions in their responses (Churchill and Iacobucci, 2005). Previous research indicates that both five point and seven point Likert-type can produce reliable and valid results (Dawes, 2008). According to Weijters *et al.* (2010), five point Likert scales can moderate the risk of missing responses, in contrast to seven or more points that could complicate choices. Embedded in the discussions above, formative measures with five point Likert-type scales are used in this study.

Many of the scales in this study are adopted from previous studies. The following paragraphs will provide a detailed explanation of the items used for each construct in this study.

Intention to share online knowledge

Ajzen (2002) measures predictive constructs in the TPB model by asking respondents' opinions on questions directly. Several items that seek to measure the construct of "behavioural intention" can be asked through "I intend to", "I try to", and "I plan to" (Ajzen, 1991). However, items should be carefully selected for different behaviours that are relevant to different populations (Ajzen, 1991).

Erden *et al.* (2012) explain that intention to share knowledge online can be defined as one's willingness to allow one's knowledge to be made available to others. There are three types of items developed to measure an individual's knowledge sharing intention: 1) general knowledge sharing intention. General knowledge is available to public; 2) explicit knowledge sharing intention; 3) tacit knowledge sharing intention (Nonaka, 1994; Grant, 1996). Explicit knowledge is defined as being codified and transferred mostly by technology, such as providing document via e-mail (Mongkolajala *et al.*, 2012). In the context of online communities, this could involve sharing opinions, articles or photos. Tacit knowledge is contrasted to the definition of explicit knowledge, in that it is not able to be written down, and is embedded in direct contact and observation of behaviour (Bock *et al.*, 2005). Tacit knowledge is revealed through practice in a particular context such as driving a car or playing a violin (Goffin *et al.*, 2011).

Against the above discussions, online knowledge involves general and explicit knowledge. Thus, items that have been employed in previous studies for measuring tacit knowledge are not suitable in this study.

Items for measuring intention to share knowledge online are summarized as below:

- I try to share knowledge with community members.
- I plan to share knowledge with community members.
- I openly share information that I gained from news, magazines and journals with other community members.
- I openly share my photo and camera related experiences with community members.

(Bock and Kim, 2002; Lin, 2007; Erden *et al.*, 2012)

Attitude to online knowledge sharing

Ajzen (2002) explains that attitude reflects an individual's evaluation toward performing the behaviour in question. The semantic differential is the most popular scaling technique in the Likert scaling procedure (Ajzen, 2002). Previous research has considered two separate components for item selection to evaluate attitudes toward intention (Ajzen, 1991, 2002; Ryu *et al.*, 2003). The first component deals with adjective pairs such as "valuable-worthless" and "beneficial- harmful". The second component often reflects a respondent's experiential quality, with bipolar adjectives such as "pleasant-unpleasant" and "enjoyable-unenjoyable". In addition, the pair "good-bad", which reflects an overall evaluation on attitude, together with the above two components are recommended for the final scale (Ajzen, 2002). "For me, sharing my knowledge with other members is: (please choose one number from 5 numbers for each line)."

1	2	3	4	5
very unpleasant				very pleasant
very unenjoyable				very enjoyable
very harmful				very beneficial

very bad

very good

very worthless

very valuable

(Ajzen, 1991).

Subjective norms

Subjective norm is defined as the sum of an individual's normative beliefs of significant others' willingness on his/her part to perform a particular behaviour, which is weighted by that individual's motivation to comply with others' willingness (Ajzen, 1991; Tylor and Todd, 1995). Normative beliefs refer to individual or collective beliefs about the prescribed / proscribed behaviours in a particular social context (Bicchieri and Muldoon, 2014). As discussed before, voluntary cooperative processing is essential to collective actions (Ostrom, 2000), thus cooperative norms are particularly related to online communities (Yen *et al.*, 2011). In this sense, subjective norms refer to perceived social pressure on the cooperative behaviours and one's motivation to be cooperative with others (Jeffries and Becker, 2008). Cooperative norms help members of an online community to understand what is expected from others and what to do in a given context, which are often inferred from text-based communications (Yen *et al.*, 2011). If members perceive an online community to be generally supportive of cooperation, they are more likely to integrate cooperative norms as their behavioural guidelines (Yen *et al.*, 2011), as a result of which, they are more likely to participate in online discussions.

Embedded in the discussions above, subjective norms in the context of online discussions can be measured through these items as follows:

- Members expend effort to maintain harmony in the forum.
- There is a high level of cooperation (e.g. replying to other members' questions and comments) among members of the online forum.
- Members are willing to sacrifice time and effort for the benefit of this online forum.

- There is a high level of sharing among members of the online forum.

(Yen *et al.*, 2011)

Perceived behavioural control

Items that are employed to measure perceived behavioural control should deal with both self-efficacy and controllability (Ajzen, 2002). Self-efficacy refers to individuals' sense of efficiency, and controllability addresses individuals' beliefs that they have control over their behaviour (Ajzen, 2002). In the case of online discussion, self-efficacy deals with members' perceptions of being able to share knowledge with others; controllability is associated with members' perceptions of ease or difficulty in sharing knowledge with others. The following represent indicators of perceived behavioural control:

- It is always possible for me to share my knowledge with network members.
- If I want, I always could share knowledge with community members.
- I feel assured that technological structures are adequate for protecting me from any problems with information systems.
- I enjoy giving my true opinion, which is not risky.

(Ryu *et al.*, 2003; Erden *et al.*, 2012)

Perceived critical mass

Critical mass refers to the tipping point after which mass actions are achieved (Oliver *et al.*, 1985). However, it can be difficult to measure such a threshold (Markus, 1990). In the context of group discussion, critical mass in different contexts is often measured through the concept of "perceived critical mass" (Shen *et al.*, 2009; Sledgianowski and Kulviwat, 2009; Lim, 2014). Items developed in prior studies are mainly designed to measure the perception of others' behaviours. For instance, Lim (2014) has developed items to understand online group buying behaviours (OGB) by directly asking for participants'

evaluation on statements such as “many people participate in OGB” and “many of my friends make comments about OGB”.

However, it is argued that what has happened before the critical point can impact on the perceived critical mass emerged after the critical point. Crossley and Ibrahim (2012) argue that there are five key claims in the theory of critical mass proposed by Oliver and Marwell (1993): 1) perception that benefits exceed contribution costs; 2) perception of the ratio of benefit to contribution cost balanced by a small group of critical mass members; 3) perception of the existence of critical mass members; 4) perception of association to a critical mass member; 5) perception of the density of communications within the online forum. Embedded in the discussions above, this study uses all five items to measure the perception of critical mass, which can be illustrated as follows in table 6:

Table 6: Measurement of perceived critical mass

Perceived benefit exceeds my cost	<ul style="list-style-type: none"> • I don't spend too much time on online discussions, but I enjoy information provided by others. • Information from my forum exceeds my knowledge.
Perceived benefit/cost balanced by critical mass members	<ul style="list-style-type: none"> • In my online forum, there are several members who give valuable suggestions because they have more resources to offer. • In my forum, there are always several members who give valuable suggestions.
Perceived existence of a giant component	<ul style="list-style-type: none"> • In my forum, only a few members are active, not many people make comments. • If those active members quit my forum, it will be a big loss.
Perception of linkage to critical mass member(s)	<ul style="list-style-type: none"> • I have friend(s) who give valuable suggestions. • I have friend(s) who are very active. • I know the member(s) who give valuable suggestions, and they become my online friend(s).
Perception of the density of online forum	<ul style="list-style-type: none"> • Many people participate in discussions. • Many of my friends participate in discussions. • Many of my friends make comments.

(Shen *et al.*, 2009; Sledgianowski and Kulviwat, 2009; Crossley and Ibrahim, 2012.)

Trust in members and trust in online forums

Trust in members is understood as interpersonal trust (Ridings *et al.*, 2002), and is proposed as one potential antecedent of attitude as discussed in the literature review. Trust in online forums indicates institutional trust (Hosmer, 1995; Mcknight and Cheveny 2002), and is assumed to positively influence attitude and perceived behavioural control as explained previously. There are numerous studies on trust, and therefore items developed in prior research can be adopted in this research, which are illustrated in Tables 7 and 8.

Table 7: *Measurement of trust in online forums*

<i>Trust in online forums</i>	
Ability	<ul style="list-style-type: none">• My forum is very competent.• My forum is able to satisfy its members.• I can expect good advice from my forum.
Benevolence	<ul style="list-style-type: none">• My forum is very concerned about the ability of people to <i>get along</i>.• If a member required help, my forum's members would do their best to help.
Integrity	<ul style="list-style-type: none">• My forum behaves in a consistent manner.• I feel fine using my forum's services since it generally fulfils its agreements.• My forum tries to be fair in dealings between members.

(Ridings *et al.*, 2002; Büttner and Göritz, 2008; Li *et al.*, 2008)

Table 8: *Measurement of trust in members*

<i>Trust in members</i>	
Ability	<ul style="list-style-type: none">• Members have appropriate skills in relation to the topics we discuss.• Members have enough knowledge about the subjects we discuss.

<i>Trust in members</i>	
	<ul style="list-style-type: none"> • Members have specialized capabilities that can add to our conversation.
Benevolence	<ul style="list-style-type: none"> • Members are very concerned about their ability to be friendly with each other • Members will not deliberately interrupt during the course of a discussion • Members will help each other solve problems.
Integrity	<ul style="list-style-type: none"> • Members throw their hearts into the communities' affairs. • Members show that they all have good morals. • Member's suggestions are the best they can offer.

(McKnight *et al.*, 1998; Ridings *et al.*, 2006)

With the constructs being identified, the following section 3.4.1.3 explains the questionnaire form, and the section 3.4.1.4 illustrates the sampling procedure.

3.4.1.3 Questionnaire form and sequence

It is noted that all items for the seven constructs are adopted from previous studies, with their wording being adapted in order to fit the context of online forums. Questionnaire forms typically use multi-item matrices. Matrices of questions can shorten a questionnaire by reducing the repetition of common initial phases for each question. An instruction is given in the beginning "please circle one for each line". Given the numbers of questions in this survey (45 in total), matrices of questions are used for this study.

However, there are three exceptions: 1) a multichotomous response form is used for the control variables such as age and education; 2) an open response form is used for asking the name of the respondent's favourite online forum(s); 3) a dichotomous response form is used for the control variable "gender". According to Churchill and Iacobucci (2005), diverse question response forms can help researchers to collect data from respondents, and may reduce monotony for responses that can further encourage questionnaire completion.

This survey is composed of three parts. After an introduction that explains the purpose of the survey, and confirms the conditions of privacy and confidentiality, the first filter question creates a skip pattern by asking whether or not participants use online forums. An answer of “no” will lead to an invitation to terminate the questionnaire. An answer of “yes” will lead to a question asking the name of favourite online forum(s) and then on to subsequent questions for completion. Churchill and Iacobucci (2005) recommend keeping the open form questions in the beginning, in order to enhance the attractiveness of the survey.

The main body of the survey asks questions about constructs as discussed above. The last section asks individual questions about personal details. Personal information is not recommended to be incorporated in the beginning of a survey, because participants may feel psychologically unsafe and stop completion (Churchill and Iacobucci , 2005).

The final version of the questionnaire was presented after the pilot test, which is included in the Appendix 1.

3.4.1.4 Sampling procedure

According to Churchill and Iacobucci (2005), four main steps are involved in a sampling procedure. These are: 1) identifying the target population; 2) identifying the sampling frame; 3) determining the sample size; 4) selecting the sampling procedure. The following sections will provide a detailed explanation of each step adopted in this study.

Target population

As discussed above, study one seeks to identify antecedents related to the social and structural influences on online voluntary contribution intention within an interest-oriented online community. The target population is therefore members of online communities, who participate in discussions around a topic within an online community, and/or view messages posted by other members.

Sampling frame

Churchill and Iacobucci (2005) argue that a random sampling technique is preferred when the population members are similar to one another in respect of important variables (such as members of online communities who are defined in terms of their knowledge sharing intention). Using random sampling, each member has an equal possibility of being chosen. In contrast to non-probability sampling that is often employed with convenience, random sampling as one type of probability sampling technique has advantage such as a higher representativeness of the sample.

Although there are difficulties of employing a random sampling technique notably due to pragmatic and situational factors such as time, finance, etc., such barriers can be less restrictive for online surveys (Wright, 2005). The cost of online surveys is minimized (Wright, 2005), because it is not necessary for the researcher(s) to be present, and the responses can be automatically downloaded into a software package which reduces not only the time of entry but also data entry error (Wright, 2005). As discussed above, a random sampling technique is considered appropriate and possible in this research.

Anderson and Gerbing (1988) argue that, when the numbers of subjects within a population tends towards infinity or it is impossible to estimate the total number (such as in the case of an online forum that do not require any formal membership or registration), the adoption of random sampling technique should satisfy two important criteria of selection: 1) each sample should be selected from the population, with equal chance of being selected. (In this research, all participants are and/or were members of online communities ;) 2) each sample should be selected independently. (In this research, responses are collected by intermediaries (zoho, soorvey, wenjuanxing) who select samples by the email address rather than by ID names, which largely reduces the overlap of sampling). In total 910 survey responses were collected. Before conducting the analysis, incomplete questionnaires were examined, leaving 910 cases. This may be due to the intermediaries only sending back the completed responses.

Sample size

Sample size is extremely important for conducting SEM analysis (Schumacker and Lomax, 2004). It is suggested that SEM requires a bigger sample size than other multivariate analyses (Hair *et al.*, 2010).

However, there is a debate on the lower boundary of sample size for SEM modelling. Bollen (1989) proposes a minimum of 5 observations per indicator. The framework presented at the end of the literature review involves 7 constructs and 42 items, which will give a required sample of 210. Another ad hoc rule of thumb suggests a 10:1 ratio of sample size to the number of free parameters (Kahai and Cooper, 2003), which leads to a sample size of 580. Free parameters are unknowns within a SEM model, which is the sum of the paths, total covariances and total variances. In this study, there are 9 paths that represent the hypothesized relationships between variables; 42 covariances, and 7 variances. Sample size should be influenced by the latent variables and their correlations rather than by the total number of observed indicators (Westland, 2010). This proposition has recently been recognised as a more robust measurement of the validity of SEM, because SEM is typically estimated in its entirety (Westland, 2010). However, this approach suggests that latent-variable SEM is asymptotic in nature, which implies an increasing number of sample sizes without boundary (Tanaka, 1987).

A problem associated with big sample sizes is the high risk of model rejection due to a lack of correspondence between model and data (Tanaka, 1987). In contrast to the big size problem, “80% of the research articles in a meta-study drew conclusions from insufficient samples” due to research cost such as time and finance (Westland, 2010, p.1). According to Veicer and Fava (1994), there is no support for rules positing a minimum sample size as a function of indicators, and the model fit is more likely to depend on the number of factor loadings and the indicators per latent variable. Embedded in the discussion above, this study collected a sample of 910 participants.

Data collection

An online survey invitation was sent to three online survey websites, namely zoho, soorvey, wenjuanxing. Wenjuanxing is a Chinese online survey website, while the others are international.

Normally, once the online questionnaire is completed, a web link is automatically generated for the survey's author. Thereafter, the author can use the snowballing technique by sending this web link to friends who may transfer that link to their acquaintances. The advantage of the snowballing technique leads to a high response rate, while the disadvantage refers that those acquaintances should be similar to each other (Cooper and Schindler, 2008). As this study follows the random survey approach, the snowballing technique is not adopted, but rather relies on the feedback from intermediaries.

Intermediaries can send one's questionnaire to others whose email addresses are registered in the intermediaries' internal database. Intermediaries can also send this survey to another author of questionnaire whose subject is different from this study. Authors of questionnaires can mutually answer each other's survey, but only once. A notification will be presented that will not allow the author to complete the same survey again. There are at least two advantages of doing so. Firstly, the response rate is enhanced; secondly, the representativeness of sampling is improved because authors may have different backgrounds and not be acquainted with each other. Although this method seems convenient for data collections, it is argued that intermediaries do not select a specified proportion from samplings, and this is not fundamentally against the basic definition of random sampling techniques.

There are limitations associated with the data collection method as mentioned above. For instance, it may be perceived as homework to answer others' questionnaires. One may complete the survey in a hurry, which may increase the risk of misunderstanding the purpose of the survey or questions themselves. In addition, online survey websites often provide incentives for completion of online questionnaire, which can increase the response rate (Deutskens *et al.*, 2004). According to Hair *et al.* (2010), incentives can reduce the quality of data because participants may want to answer the questionnaire more than once.

However, the risk of multiple responses can be reduced as online survey websites can distinguish IP addresses and contact email addresses. As discussed above, a very high response rate for this research (100%) is generated, because intermediaries may only send the completed responses by deleting partially completed ones. Hence, this does not demand for the analysis of missing data that can either be randomly missing or system missing.

Sampling errors and no sampling errors

According to Churchill and Iacobucci (2005, p. 679), sampling errors suggest the differences between the observed value and the true value related to a variable. Sampling errors are inevitable because there is always variation between the observed value and the true value in measurement (Hair *et al.*, 2010). However, sampling errors can be moderated by increasing the sample size (Hair *et al.*, 2010).

Hair *et al.* (2010) explain that non-sampling errors can be classified into two categories, which are non-observation errors and observation errors. Non-observation errors often suggest non-response errors. In addition, they occur when responses differ to those who respond (Armstrong and Overton, 1977). In this study, non-observation errors are minimized because there is no blank response in the answer sheet. Observation errors involve administration techniques, measurement scales, editing and coding, and even analysis of the data, which is more difficult to manage in the data analysis process (Churchill and Iacobucci, 2005). The risk of observation errors can be reduced by adopting scales and measurements validated in previous research, followed by a pilot test (Hair *et al.*, 2010). In this study, the survey questionnaire is additionally firstly pre-tested in an effort to try to reduce the risk of misunderstanding the survey instrument by respondents. This process, together with the pilot test, will be further explained in the following sections.

3.4.2 Pilot test

A pre-test is recommended before the pilot test (Churchill and Iacobucci, 2005; Hair *et al.*, 2010). The pre-test helps the researcher to confirm whether the questionnaire is tidy, and to estimate the approximate

completion time for respondents (e.g. Hair *et al.*, 2010). The pilot test provides the researcher with an overview of the methodological and analytic techniques of the study (e.g. Churchill and Iacobucci, 2005).

The pre-test is conducted with a sample of ten acquaintances of the researcher, comprising 5 Chinese, 4 French and 1 British person. The Chinese and French version of the questionnaire is translated from the English version. A native French speaker is asked to verify the final version in French. Respondents used a web link to the survey and commented on the layout and the appearance of the survey. The pre-test automatically recorded the time of completion in total and for each stage, and this resulted in an average of fifteen minutes being taken to complete the survey.

A few points are noted during the pre-test which were subsequently incorporated into the pilot test:

1. It appeared that respondents use different browsers and operating systems which include answering through mobile phone (4), through Firefox (4) and through Internet explorer (2), inferring that people have different habits of surfing online.
2. It was recommended to write the questions in bold with size 14, and the answers in a lighter font with size 12 in order to create a clear navigation path.
3. It was recommended to use colour shading to attract attention to group questions. The light grey colour shading is used in alternate groups of questions.

The pilot test is conducted on a sample of 80 participants, which gives a potential indication of the reliability of the measurement scales adopted in this study. The procedure for the pilot test is the same for performing the main test, and explained in the following sections. In summary, the results indicate the multi-item scales are acceptable, with results presented in chapter 4.

3.4.3 Preliminary data analysis

Hair *et al.* (2010) recommend undertaking preliminary analysis on the data before conducting SEM, because this step can increase the reliability and validity of final data analysis by allowing cleaning and screening of any problems with the data set (Tabachnick and Fidell, 2007).

One advantage of using an online survey is that the results can be downloaded directly into Excel and thereafter into the statistic package SPSS, which reduces the risk of data entry errors. Data is firstly coded and labelled using syntax in SPSS whereby values for all items are allocated within the same scales. Each variable name is labelled with a few words, and variable values are coded from “1” representing “strongly disagree” to “5” representing “strongly agree” for all variables, except for the control variables. Sex is coded “1” (male) and “2” (female). Age is coded from “1” to “4” meaning “<20”, “21-35”, “36-50” and “>50”. Finally the variable “Education” is coded starting “1” (high school) to “5” (PhD) sequentially. After this step, the data set is ready for analysis.

3.4.3.1 Missing data and outliers

Both missing and outlying data can adversely influence the results generated by SEM (Hair *et al.*, 2010). As mentioned above, missing data is not an issue in this study because the online survey only allowed for the submission of fully completed questionnaires.

SPSS V22 can detect outlying responses (outliers) by examining the distribution of z scores (standard scores) for all variables. For sample size greater than 200, which is the case of the main study, outliers are cases with z-scores beyond ± 4 (Hair *et al.*, 2010). In SPSS, z-score for each value is calculated through:

$z - scores = \frac{(\bar{x} - x_i)}{\sigma}$ (3.1), where \bar{x} is the mean values of x_i which represents i^{th} value in the data set. σ is

$$\sigma = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N}} \quad (3.2), \text{ with } N \text{ being the sample size.}$$

the standard deviation which equals to :

Each individual outlier should be examined carefully, and it can be deleted only if it is not representative of the population (Hair *et al.*, 2010). In this study, the search function in SPSS is used in order to identify outliers, with the result that all values are within the acceptable range.

With no issues in terms of missing data and outliers to address, the next step is to check whether data is normally distributed because it is an important assumption for the further statistical tests which will be illustrated in the following sections.

3.4.3.2 Descriptive analysis

The descriptive analysis should be performed after further cleaning and screening (Hair *et al.*, 2010). Results generated from the descriptive analysis include the mean for each individual variable as well as the overall score for the multi-items scales. The descriptive analysis is a simple summary of data features in terms of data frequency and distribution. However, it can provide a quick insight of the key features of the study data, for instance, by gaining a first estimation of the central tendency by looking at means in the frequency table.

Testing the normal distribution of variables is very important as many statistics analyses are embedded in the assumption of normal distribution (Baumgartner and Homburg, 1996). SPSS V22 initially provides the Kolmogorov-Smirnov statistic with a significant value to test for non-normality of data, which shows that data is not normally distributed ($p < 0.001$, two- tailed). However, the Kolmogorov-Smirnov statistic test is sensitive to a large sample size, because a large sample size can statistically support the deviation against means even when such variances are meaningless and negligible (Tabachnick and Fidell, 2007).

For 910 sample size in this study, the kurtosis and skewness tests are further employed in order to check the data distribution. Kurtosis examines the shape of data distribution, while skewness can tell the position of distribution (central, left or right) compared to the normal distribution (Hair *et al.*, 2010). In

SPSS, the formulas for calculating skewness and kurtosis are illustrated as follows:

$$Skewness = \frac{\sum z^3}{N}$$

(3.3); $Kurtosis = \left(\frac{\sum z^4}{N} \right) - 3$ (3.4). In these two formulas, z is the z-scores that measure the differences between the raw values and their means; N is the sample size.

The acceptable ranges of normality for the skewness and kurtosis values are ± 2 (George and Mallery, 2010), with results applicable for the normality assumption in this study. Although the Kolmogorov-Smirnov statistic is not satisfied for the normality test, it is argued that the normal distribution of data rarely occurs in practice in social science research and deleting non-normal distribution variables often leads to the dropping of important information that can otherwise be obtained from the original data set (Hair *et al.*, 2010). Hair *et al.* (2010) further argue that the maximum likelihood estimation (MLE) technique adopted by SEM modelling can be a robust measurement even with non-normal distribution values. Furthermore, a sample size greater than 200 will minimise the effect of non-normal distribution (Diamantopolous *et al.*, 2000). Thus, the data is considered normally distributed because the values for skewness and kurtosis are within the acceptable range.

3.4.3.3 Response rate and non-response analysis

Sampling errors may occur because the whole target population is not included in the survey responses (Churchill and Iacobucci, 2005). Although it is not possible to accurately measure non-responses, it is suggested that non-responders are likely to answer in a similar way as late responders (Armstrong and Overton, 1977). In this study, there are two waves of responders: those who respond during the first 2 weeks and those who respond later. Independent sample t-tests to compare the mean responses for each construct between early and late responders were used. Similar sample characteristics can suggest a valid sample in the context under study (Malhorta *et al.*, 1996).

The t-tests were performed using the software SPSS V 22. The t-value was calculated using the

formula: $t - values = \frac{\bar{x}_1 - \bar{x}_2}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ (3.5), where $s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$ (3.6) representing the pooled

standard deviation that assumes the two independent samples have equal variance; s_1, s_2 indicating the standard error of the means from the two different groups \bar{x}_1 and \bar{x}_2 ; n_1 and n_2 are the sample size for the earlier responders and later responders respectively. The standard error refers to the estimation of the standard deviation of the sampling distribution which is often measured through the differences between means that is a consequence of sample error (Hair *et al.*, 1998). A significant t-statistic of <0.05 that is the result of comparison between t-value and the critical t-value with the degree of freedom $(n_1 + n_2 - 2)$ will accept the alternative hypothesis of significant differences between groups $(\bar{x}_1 \neq \bar{x}_2)$ (Hair *et al.*, 2014).

SPSS generates the t-test significance statistic and the Levene's test simultaneously. The Levene's test is to examine whether equal variance between the groups can be assumed. If equal variance between the groups is expected, the variability within the responses can be assumed to have occurred at random (Hair *et al.*, 2014). In Levene's test, the F-max is compared with the critical value $F_{\max crit}$ with degree of freedom $(n_1 - 1, n_2 - 1)$ which will generate a significance value. F-values is to compare the larger

variances to the smaller variances:
$$F_{\max} = \frac{\max(s_1^2, s_2^2)}{\min(s_1^2, s_2^2)} \quad (3.7).$$
 If Levene's statistics are significantly smaller than 0.05, the null hypothesis of equality of variances should be rejected (Hair *et al.*, 2014).

The effect size of any significant difference should be calculated before deciding whether the significance statistic can be generalized for the whole target population (Hair *et al.*, 2010). The effect size for t-test is

calculated through the formula: effect size for t-value = $\frac{t^2}{t^2 + (N - 1)}$ (3.8), where t represents the t-value, and N is the sample size. Cohen's (1988:284-287) criterion classifies the effect size into three categories: 1) any effect size values between 0.01 and 0.05 is considered small; 2) any effect size values 0.06 and 0.13 is medium; 3) any effect size values greater than 0.14 is big. In addition, the effect size value can be multiplied by 100 for obtaining the percentage of variability in responses (Hair *et al.*, 2010). This test can further suggest the level of variability within samples, which is one consequence of non-response error. In

this study, results suggest that the equal variances between the earlier responders and later responders can be assumed, and that results from this study could be generalized.

3.4.3.4 One-way ANOVA

Tabachnick and Fidell(2007) explain that it is necessary to test whether the differences in characteristics of groups occur across the data set. In addition to the comparison between the earlier and later responders mentioned above, this study examines whether significant variances exist in terms of control variables namely gender, education level and age. A result of equality of variances between groups can further provide evidence that the results produced through SEM could be generalised (Tabachnick and Fidell, 2007).

The same independent t-test as explained above is performed to compare two independent samples e.g. females and males. To compare more than two groups' means simultaneously, analysis of variance (ANOVA) techniques are employed for the variable education and age. The objective of a one-way ANOVA test in this study is to compare the calculated F-value (which is different from the F-value in Levene's test in the independent t-test) to the $F_{critical}$ value in the table for the F distribution at p=0.05 level of significance. The null hypothesis of equal variances between groups should be rejected if the calculated F-values is greater than the $F_{critical}$ with degree of freedom (numbers of groups -1, numbers of groups*(numbers of observations within group -1)).

The formulae employed to calculate the F-values are illustrated as follows:

$F = \frac{MS_B}{MS_W}$ (3.9), with: MS_B representing the between-groups variance (chance variance + treatment

effect) : $MS_B = \frac{\sum_{j=1}^k (\bar{x}_{.j} - \bar{\bar{x}})^2}{k-1}$ (3.10), whereby, $\bar{\bar{x}}$ is the grand means $\bar{\bar{x}} = \frac{\sum_i \sum_j x_{ij}}{n_i * j}$ (3.11); $\bar{x}_{.j}$ is the means of

the group j; And MS_w denoting the within-group variance (chance variance):
$$MS_w = \frac{\sum_{i=1}^{n_j} (x_{ij} - \bar{x}_j)^2}{n_i * j - k} \quad (3.12),$$

whereby x_{ij} is the i^{th} observed value in the group j; \bar{x}_j is the sample means of the group j; and K represents the numbers of groups.

The effect size of ANOVA is calculated to assess the extent of differences by the formula:

Effect size of ANOVA = $\frac{MS_B}{MS_B + MS_w}$ (3.13). The same criteria as for the independent t-test are applied on the effect size in ANOVA. If a big effect size of ANOVA is calculated, it is necessary to perform multi-group analysis during the SEM process which requires different models embedded in the different groups of samples (Bryne, 2013) If the assumption of equal variability between groups is satisfied, one structural equation model can be developed for the whole sample. In this study, the results indicate a small effect size and one structural model is satisfactory for the whole data set. The following sections will discuss the structural equation modelling process.

3.4.4 Structural equation modelling

3.4.4.1 Introduction to SEM

Structural equation modelling (SEM) is the main method used for testing the hypotheses illustrated in chapter 2. SEM has been recognised as the second generation of multivariate techniques, and has become popular within marketing research (Anderson and Gerbing, 1988; Bagozzi, 2010).

Processing SEM needs the development of two sub-models, a measurement and a structural model (Anderson and Gerbing, 1988; Schumacker and Lomax, 2004; Bryne, 2013). The measurement model is

composed of two equations, which can be expressed as followings:
$$\begin{matrix} X \\ (qx1) \end{matrix} = \begin{matrix} \wedge_X \\ (qxm)(nx1) \end{matrix} \xi + \begin{matrix} \delta \\ (qx1) \end{matrix}$$

$$\begin{matrix} Y \\ (px1) \end{matrix} = \begin{matrix} \wedge_Y \\ (pxm)(mx1) \end{matrix} \eta + \begin{matrix} \varepsilon \\ (px1) \end{matrix} \quad (3.14),$$
 where X refers to (qx1 vector) observed exogenous variables, which is measured through the factor loadings of the

observed x variables (\hat{x} , $q \times n$ vector), the common factor ξ ($n \times 1$ vector) that may have direct influence on more than one observed variable, and the error terms ε for the observed indicators ($q \times 1$ vector). The endogenous variable Y ($p \times 1$ vector) is measured through \hat{y} representing the coefficient matrix ($p \times m$) of the relationship between observed indicators toward the latent variables η and the residual vector ε ($p \times 1$). The structural model explains the prediction for an endogenous dependant in a

path model, which is illustrated as:
$$\underset{(mx1)}{\eta} = \underset{(mxm)(mx1)}{B_{\eta}} + \underset{(mxn)(nx1)}{\Gamma}\xi + \underset{(mx1)}{\zeta} \quad (3.15),$$
 where η is the endogenous construct, ξ is the exogenous construct, and B is the coefficient matrix ($m \times m$) of relationships between endogenous variables; Γ is the matrix ($m \times n$) of relationships between exogenous ($n \times 1$) and endogenous variables; ζ refers to the column of error term ($m \times 1$) for endogenous variables.

The measurement model involves the confirmatory factor analysis (CFA) technique which examines how the latent variables are measured by indicators (Weston and Weston and Gore Jr, 2006; Hair *et al.*, 2010). The structural model indicates the coefficients between latent variables (Anderson and Gerbing, 1988; Weston and Gore Jr, 2006). That is, the measurement model seeks to reduce a set of observed variables into manageable constructs by examining the inter-correlation between items within the latent variable; the structural model is to test the hypothesised regression relationship between latent constructs.

SEM is a more powerful analysis technique than multiple regressions, because all path analyses can be included in one model, while the multiple regressions run separate examinations of each hypothesized path (Gefen and Straub, 2000). Another benefit of SEM is that it allows both observed and latent variables to be analysed at once (Schumaker and Lomax, 2004). In addition, SEM can generate reliable results as it takes into account the measurement errors within the observed variables (Steenkamp and Baumgartner, 2000; Diamantopolous *et al.*, 2000; Schumaker and Lomax, 2004). Finally, SEM can reduce the risk of multicollinearity by creating the measurement model which allows the inter-correlated observed variables to be classified into latent variables (factors) (Iacobucci, 2009; Verbeke and Bagozzi, 2000; Lee and Hooley, 2005).

In this study, multi-item scales are developed for constructs as discussed above. If using regression analysis, each regression relationship between items and variable should be separately examined, which could create a high multi-collinearity between variables. It is therefore suggested that SEM is used for this study. According to Schumaker and Lomax (2004), SEM uses sophisticated quantitative analysis methods to confirm or disconfirm theories, resulting in more reliable and useful outputs.

There are two classifications of SEM-based analysis models, which are covariance based SEM (CB-SEM) and partial least squares based SEM (PLS-SEM) (Hair *et al.*, 2014). In contrast to PLS-SEM, CB-SEM seeks to minimise the differences between the tested and estimated covariance matrices (Hair *et al.*, 2014). Hair *et al.* (2014) argue that PLS-SEM is preferred for the exploratory works, as it can be conducted on a small sample size. CB-SEM technique is widely applied in the marketing research (Hair *et al.*, 2010; Hair *et al.*, 2014) and information systems literatures (Wright *et al.*, 2012), because it is helpful in confirmatory factor analyses (Byrne, 2010; Hair *et al.*, 2010; Hair *et al.*, 2014) and evaluating models incorporating multidimensional constructs (Wright *et al.* 2012). However, the limitations associated with CB-SEM should not be ignored. It requires normally distributed data as well as 5-10 indicators loading toward each latent variable at first (Hair *et al.*, 2014). In addition, if it happens that less than three indicators remain loading toward each factor after the convergent analysis, this is a challenge to meet the requirements for at least three indicators to predict each construct in the structural model (Hair *et al.*, 2014). In this study, the multidimensional constructs such as “trust in online communities”, “trust in members” and “perceived critical mass” are measured through 5 observed variables after the EFA stage. In addition, data is considered normally distributed as the results generated by the asymmetry and kurtosis tests are within the acceptable ranges for normality. As such, the CB-SEM is performed.

3.4.4.2 CB-SEM processing

As mentioned above, structural equation modelling comprises a measurement model and a structural model. The measurement model seeks to identify the regression relationships between indicators and latent constructs (factors), and the structural model aims to suggest the relationships between latent

constructs. According to Anderson and Gerbing (1988), the measurement model can be developed and evaluated with respect to its measurement reliability and validity before producing the structural model. The alternative one-step approach suggests examining simultaneously both the measurement and structural models. However, the one-step structural modelling is criticised, as it may generate a mis-specified model (Anderson and Gerbing, 1988; Gallagher *et al.*, 2008).

Construct validity and reliability seeks to assess the extent to which the measure of a construct is a true reflection of the construct, and can be assessed with respect to face validity, convergent validity, discriminant validity and nomological validity (O'Leary-Kelly and Vokurka, 1998). Construct validity is important because it assesses whether the observed variables are inter-correlated with observed variables loading onto different constructs (Bryne, 2013; Bagozzi, 2010).

Face validity refers to the examination of whether the constructs and the items that load on them have both theoretical and practical sense (Hair *et al.*, 2010). Exploratory factor analysis (EFA) is able to assess the face validity. Convergent validity examines the internal consistency of the constructs through measuring the degree to which the observed variables loading onto a factor have similar meaning (Hair *et al.*, 2010). Discriminant validity assesses whether each construct will not be the reflection of other constructs, and is unique within the model (Churchill, 1979). Nomological validity refers to examination of the inter-construct correlation (IC) which should be theoretically meaningful (Hair *et al.*, 2010).

Gerbing and Hamilton (1996) explain that the cross-validation technique can be applied for increasing the reliability of the measurement model. In this study, the whole sample is separated randomly into two groups, a training dataset (around 50% of all responses) and a testing dataset. Principal component analysis (PCA) is conducted on the training dataset and results are cross-validated with the testing dataset where the confirmatory factor analysis (CFA) is performed. The whole sample is used for examining the final measurement model and for establishing the structural model. EFA seeks to assess the face validity that verifies whether or not the constructs and the items that load on them have both theoretical and practical sense (Hair *et al.*, 2010), and PCA as the tool of EFA is the factor-reducing technique without

losing too much information. CFA is relatively more complex which is able to test the convergent, discriminant and nomological validity of factors (latent constructs) and the overall model fit (Anderson and Gerbing, 1988). The following sections will discuss CFA techniques using CB-SEM.

Preparing for CFA

Exploratory factor analysis (EFA) seeks to group indicators (manifested/observed variables) into factors based on their similarity (Hair *et al.*, 2010). As its name suggests, EFA is completely exploratory without either specifying the numbers of factors to be extracted from the dataset or indicating which indicators belong to factors. EFA is therefore conducted to understand the reliability of multi-scales employed in this study.

The Kaiser-Meyer-Olkin measure of sampling adequacy (MSA) and Bartlett's test of sphericity are performed before assessing the structure of factors generated from the analysis. Kaiser's criterion is used in this study as it is completely exploratory without making assumptions on the numbers of factors to be extracted from the dataset (Hurley *et al.*, 1997; Yeomans and Golder, 1982; Lee and Hooley, 2005). Kaiser's criterion explains that factors remain in the analysis only if their eigenvalue is greater than or equal to one (Hair *et al.*, 2010). MSA tests whether the sample size is sufficient for factor analysis, whose value should be greater than or equal to 0.5. A MSA value over than 0.9 and close to 1 is preferred.

Bartlett's test tells whether the original correlation matrix is identical. A Person's correlation coefficient value over 0.3 suggests that two items are correlated but greater than 0.9 will lead to the deletion of one item due to multicollinearity in the data. A significant Bartlett's test ($p < 0.001$) rejects the null hypothesis that the correlation matrix is identical, and suggests that there exist inter-correlations within variables that often happens, which is the case with this study.

Principle components analysis (PCA) is the extraction technique used in this study, as it is assumed that there is no unique variance within the observed variables. In fact, PCA is the most commonly used technique in factor analysis (Hair *et al.*, 2010). The rotation methods applied within PCA can be either

orthogonal method that assumes factors in analysis being uncorrelated or oblique rotation method that assumes the factors in analysis being correlated (Gorsuch, 1990). Tabachnick and Fidell (2007) recommends that the oblique rotation method can be conducted firstly, checking the factor correlation matrix for correlation around 0.32 or above, which is the boundary for warning the orthogonal method due to at least 10% of variances overlapped among factors. In this study, results suggest that the oblique rotation method is preferred over the orthogonal method. SPSS 22 offers the direct oblimin and promax methods for the oblique rotation, and both methods generate similar results (Brown, 2009). The direct oblimin method is therefore performed as it is the most commonly employed one for the oblique rotation (Kim and Mueller, 1978).

PCA calculates the eigenvalues of the correlation matrix. An eigenvector is considered important if its value is greater than 1. The total variance explained by the eigenvectors should exceed 50% (Hair *et al.*, 2010). The communality and component matrix outputs of PCA assesses whether all observed variables loading toward a factor have a significant level of over 0.05 (two tailed test), by examining communalities and factor loadings that should be equal to or greater than 0.50 (Hair *et al.*, 2010). The communality explains the common variance shared within a variable. Before extraction, PCA assumes that all variances explained by variables are common. While with possible information loss during the extraction of variables, the remaining variables can only explain some variances of the data. Communality values of 0.5 imply that 50% of variance is explained by the variable (factor). The factor loadings explain how much the observed variables are correlated with the component. PCA assumes that the observed variables are linearly combined with components, and a weighted factor loading value of 0.5 shows that the observed variable determines 50% of variances explained by the components. However, it is noted that three or more observed indicators for each latent variable are recommended in SEM analysis (Hair *et al.*, 2010). That is, indicators that do not meet the two requirements as mentioned above are deleted sequentially with respect to their factor loading and communalities starting from the lowest values in order to keep at least three observed variables to predict one component after PCA. One main criticism

of PCA is its limitation in reducing observed variables loading toward the latent variable (Hair *et al.*, 2010), and this limitation will be compromised with CFA analysis.

Cronbach's alpha (α) is a commonly used technique within EFA for assessing internal consistency within measurement scales, and can be used following factor analysis (Cortina, 1993). The values generated from Cronbach's alpha (α) test range from 0 to 1, where a value of 1 suggests perfect consistency between indicators and constructs. The recommended acceptable level of Cronbach's alpha (α) is over than 0.7 (Cronbach, 1951; Nunnally, 1967; Cortina, 1993). The confirmatory factor analysis is conducted following PCA.

Confirmatory factor analysis using CB-SEM

This section explains how the measurement model is evaluated using CB-SEM. As mentioned above, this study involves three multidimensional constructs, which are "trust in members", "trust in online communities" and "perceived critical mass". Wright *et al.* (2012) recommend creating seven models to assess the multidimensional constructs validity with respect to their convergent validity, discriminant validity and reliability, which is explained in the following table 9.

The first model hypothesizes that all observed indicators are loading toward one first order factor. Bad model fit indices may suggest that a factor is multidimensional as indicators cannot be all loading toward one factor. Otherwise, the second order factor may not be well formed.

Table 9: An illustration of CB-SEM process

Step 1	Step2	Step3	Step4
Run model 1: first order factor model	Run model 2: freely correlated first-order factors	Run model 3: tests of discriminant validity	(a) Run model 4: parallel Model
Evaluate fit statistics: fit should be poor	(a) Evaluate fit statistics: improved fit over model 1 supports dimensionality	(a) Run two freely correlated factors then constrain the	(b) Run model 5: tau equivalent Model
			(c) Run model 6: congeneric Model

Step 1	Step2	Step3	Step4
	(b) Evaluate factor loadings: significant loadings support convergent validity	correlation (b) Evaluate fit statistics: significant χ^2 change supports discriminant validity (c) Repeat for each pair of first-order factors	Compare models 4, 5, and 6 to select the best fitting model.

Source: Wright *et al.* (2012)

The second model allows 15 predictors loading on their assumed factors that are three freely correlated first order factors: “trust in members”, “trust in online community” and “perceived critical mass”. If the overall model fit indices are improved comparing them with that generated from model one, there is some evidence that constructs are multidimensional. The convergent validity is evaluated by examining the parameters that should be significant at the level of .001 (two tailed) (Tabachnick and Fidell, 2007). In addition, it is necessary to check the standardized factor loadings which should be greater than or equal to 0.5 as the recommended levels, or greater than 0.7 as the preferred levels (Hair *et al.*, 2010). For this reason, all observed variables with standardised loading lower than 0.5 are excluded.

The third model seeks to test the discriminant validity between paired factors. It is noted that the other four non-multidimensional factors which are “attitude”, “perceived behavioural control”, “subjective norms” and “intention to contribute knowledge online” are included in the discriminant validity analysis. The discriminant validity is supported if there are significant changes between the chi-squared values. In Amos, the Chi-squared tests are performed through two models representing with and without correlated factors (Zait and Berteau, 2011).

The parallel, tau and congeneric models involve the second order factors which are the concepts to be measured. The parallel model treats each dimension equally by restricting both the factor loadings and

residual variances. The tau model releases the residual variances but restricts the equal loadings, thereby allowing an examination of the internal consistency. The congeneric model frees all constraints in the parallel model, which is able to measure construct reliability. Comparing the model fit indices generated from the parallel, tau and congeneric models, the best fit model is considered as the reliability estimations. The second order factor model can assess the nomological validity because the first order factors are correlated due to the second order factor (Liu *et al.*, 2012). The nomological validity is often conducted within the nomological network where the multidimensional constructs have both antecedent factors as well as the consequent factors, which is not the case of this study. The multidimensional constructs trust and perceived critical mass are the causes of intention to contribute knowledge online mediated by the attitude, perceived behavioural control and subjective norms. They are correlated mainly due to the same mediation factor subjective norms.

The score of an observed variable X_i is equal to the sum of the true scores of the observed variable and its variances, which can be denoted as : $X_{ip} = T_{ip} + \varepsilon_{ip}$ (3.16). That is, there are p numbers of the observed measures for the variable X_i , which is determined by the total p numbers of true scores for X_i and the sum of variances associated with p^{th} observe. If there are k numbers of observed variable X_i (1, 2,...k) that are meant to measure a composite variable Y that is created in SEM by adding the error term to each individual observed variable. As well as taking into account the variances shared between observed variables (Graham, 2006), Y is the direct results of the sum of the total k numbers of the true scores for X_i and the total variances of each observed item.

In the parallel model, the true scores for all predictors are constrained to be equal; likewise, the variances in the observed variable scores are the same for all observed variables. Therefore, each item i for individual p can be presented as: $Y_{ip} = T_p + \varepsilon_p$ (3.17) (Graham, 2006). The parallel model indicates that the same measures for a composite variable on the same scale can be applied on multiple occasions where the true scores will not change and that the variation of each measure will be the same.

With the tau model, it is assumed that the true scores for all items are the same, while each item error term is not set to be equal. That is, each item i for individual p can be shown as: $Y_{ip} = T_p + \varepsilon_{ip}$ (3.18). The tau model implies that although the item's true scores being measured on the same scale are the same, differences of the degree of precision for each item (different means) are expected (Graham, 2006). For example, the two questions adopted in this study to measure communication wideness (the construct of "perceived critical mass") are "many of my friends participate in discussions" and "many of my friends make comments", which is measured on the same five point Likert-scale from "strongly disagree" to "strongly agree". Although responses to these two questions that are designed to measure the same dimension have a similar distribution, the means of these two questions are slightly different at 3.54 and 3.51 respectively. This may be one consequence of the different precision demonstrated by these two questions.

The congeneric model is often chosen as the model for testing reliability (Hair *et al.*, 2014). It assumes linear relationships between items' true score, which can be denoted as: $Y_{ip} = [a_k + \beta_k(T_{ip})] + \varepsilon_{ip}$ (3.19). That is, it allows the true scores for each item to differ, as well as the degree of items' precision and variances across items (Graham, 2006). Having compared the model fit indices for the parallel, tau and congeneric models, the congeneric model demonstrates the best model fit.

Table 10: Summarizes the conditions for construct validity

Face validity	Results from EFA are confirmed for both theoretical and practical meaning of constructs.
Convergent validity	1. Standardised loadings significant at the 0.001 level (two-tailed) 2. Standardised loadings >0.50 or preferred >0.70
Discriminant validity	Chi-square difference test
Reliability	Choosing the best model fit from the parallel, tau and congeneric models

The structural model

The next and final step after examining the latent variables within the measurement model is to estimate the relationships between constructs in the structural model (Anderson and Gerbing, 1988). As discussed previously, the structural model looks at “the nature and magnitude of the relationships between constructs” (Hair *et al.*, 2010, p. 729).

Paths that represent the hypothesised relationships between latent variables are added into the measurement model. The overall model fit indices are firstly assessed with the choice of model indices explained in the following section. The significance of path values is determined by t-values, and the greater the significance, the more likely that the relationships between latent variables are valid (Tabachnick and Fidell, 2007).

Standardised estimates for paths are interpreted as their regression values (Weston and Gore, 2006). The overall variance R^2 that explains the covariance between latent variables is obtained through squaring the errors associated with independent latent variables (e.g. D^2), and subtracting them from 1 (Weston and Gore, 2006) : $1 - D^2 = R^2$ (3.20). R^2 can be interpreted according to the recommendations of Hair *et al.* (2010):

- 0.10-0.29 = reasonable predictor of outcome construct
- 0.30-0.49 = good predictor of outcome construct
- 0.50-0.69 = very good predictor of outcome construct
- 0.70+ = extremely good or potentially suspicious model

Testing an alternative model

A positive degree of freedom suggests an over-identified model. An over-identified model means that there is more than one possible solution to explain the data. An alternative model should be taken into account (Homburg, 1996; Bryne, 2013). However, it is noted that there should be either theoretical or

empirical support for the additional paths in an alternative model (Schumaker and Lomax, 2004; Hair *et al.*, 2010).

There are three ways for comparing the alternative model and the estimated model (Weston and Gore, 2006): 1) examining the path significances (the significant parameters estimated in the alternative model suggest that the alternative model provides an explanation of the sample data); 2) considering the changes in explained variances (it is expected that the alternative model will generate an increase or no change in explained variances); and 3) test mediation. The following section will explain the model fit indices adopted in this study.

3.4.4.3 Testing of model fit

As discussed above, both the measurement model and the structural model involve the comparison of model fit indices. It is therefore important to discuss the reason why such a model fit indices are chosen for analysis. The indices can be classified into two categories, goodness of fit (GOF) and badness of fit (BOF). Commonly used GOF indices include parsimonious fit indices, absolute fit indices, and incremental fit indices (Tabachnick and Fidell, 2007; Hair *et al.*, 2010). It is suggested that at least one from the three types of model fit indices should be assessed (Hair *et al.*, 2010). It is also recommended to evaluate both GOF as well as badness of fit (BOF) (Bryne, 2013).

The basic logic of the chi-square (χ^2) test is to assess whether the observed results are equal to the

expected results, and can be shown as:
$$\chi^2 = \sum_{i=1}^N \frac{(\text{observed}_i - \text{expected}_i)^2}{\text{expected}_i} \quad (3.21)$$
, where N is the possible

outcomes. That is, χ^2 becomes bigger as long as the differences between the observed and expected (implied) results are larger. In SEM, chi-square tests tend to measure whether the observed covariance matrix, Σ , is equal to the expected covariance matrix, $\Sigma^{(\theta)}$ (Hoyle, 2014). A significant χ^2 at 0.05 levels ($p < 0.05$) indicates that the null hypothesis of equivalency of the covariance matrixes should be rejected (Hooper *et al.*, 2008). For discriminant validity within the measurement model, significant chi-squared

differences between the two models comprising paired latent factors with and without correlation between them are expected. In the badness of fit test with the structural model, the researcher is interested in obtaining a non-significant χ^2 , because it indicates that the implied model can significantly reproduce the same variance-covariance relationships in the matrix (Schumacker and Lomax, 2004). However, the chi-square test is sensitive to size. With a larger sample size (above 400), χ^2 tends to report a significant level (Schumacker and Lomax, 2004), and almost immediately reject the estimated model (Bagozzi, 1988; Bentler and Raykov, 2000). To minimise the impact of the sample size on the structural model, the relative/normed chi-square (χ^2/df) should be within the acceptable range between 2 and 5 (Hooper *et al.*, 2008). As in this study, the sample size is larger than 400, the normed chi-square is referred as a badness of fit statistic.

The root mean square error of approximation (RMSEA) is adopted in this study as an alternative measure to Chi-square (χ^2) (Bentler, 2007). In contrast to chi-square (χ^2), a small sample size can lead to a model being rejected (Tabachnick and Fidell, 2007). RMSEA as one badness of fit indicator is able to estimate even a complex model (Baumgartner and Homburg, 1996). RMSEA is calculated through (Hoyle, 2014) :

$$RMSEA = \sqrt{\frac{\max(\chi^2 - df, 0)}{df(N-1)}} \quad (3.22), \text{ where } df \text{ represents the degree of freedom, } N \text{ the sample size, and } \chi^2$$

Chi-square. The degree of freedom equals to $k(k-1)/2$, where k is the numbers of variables. A value of RMSEA between 0.05 and 0.07 indicates a good model fit (Hooper *et al.*, 2008). The lower and upper boundary of 90% confidence interval and the probability of significant RMSEA (< 0.05) can be shown in the statistical software Amos.

Goodness of fit indices (GFI) is another alternative to the chi-square test (Hooper *et al.*, 2008). GFI is

calculating through: $GFI = 1 - \frac{e'We}{s'Ws}$ (3.23) (Bentler, 1983), where $e'We$ represents the weighted sum of squared residuals from a covariance matrix, and $s'Ws$ is the weighted sum of squared covariance and variance (Hoyle, 2014). GFI varies between 0 and 1, and the general cut-off is at least 0.9. However, GFI

is sensitive to the sample size, df and the numbers of parameters within the model. With increasing sample size and number of parameters, GFI is more likely to be larger, while high degrees of freedom will have a downward effect (Cooper *et al.*, 2008). As the sample size in this study is big (460 for CFA and 910 for Structural model), GFI, which is often reported, is not reported here.

Incremental fit indices are one of most useful model fit indices (Widaman and Thompson, 2003) and measure how well the estimated model fits compared with a null model (Hair *et al.*, 2010). The

comparative fit index (CFI), which is calculated through:

$$CFI = 1 - \frac{\max(\chi_M^2 - df, 0)}{\max(\chi_B^2 - df, 0)} \quad (3.24)$$

developed by Bentler (1990) is the incremental fit indice assessed in this study. CFI assumes that all observed variables loading toward a factor are not correlated, and compares the measured sample covariance matrix with its null model, also as the baseline model (Cooper *et al.*, 2008). CFI is reported as it is the fit indices least affected by the sample size (Fan *et al.*, 1999). According to Bentler (1990), the values of the incremental fit indices range from 0 to 1, and a value that is closer to 1 (≥ 0.9) is normally accepted as the cut-off for acceptable goodness of fit.

Parsimonious fit measures consider the numbers of parameters in the estimated model (Mulaik *et al.*, 1989). According to Baumgartner and Homburg (1996), they are particularly useful goodness of fit indices for comparing alternative models with different levels of complexity. The parsimony goodness of fit index (PGFI) is one important parsimonious fit measure, and is based on GFI by adjusting the loss of

degree of freedom (Cooper *et al.*, 2008); it is adopted in this study:

$$PGFI = \frac{df_M}{df_B} GFI \quad (3.25)$$

As PGFI is penalised in a complex model, its value is reduced compared with GFI. According to Mulaik *et al.* (1989), a value of PGFI that is greater than 0.05, often indicates a good model fit.

Table 11: Summary of the fit indices adopted in this study

Absolute fit index:	
RMSEA;	<0.07 suggests a good fit.
χ^2/d	Accepted range between 2 and 5.
χ^2 differences tests for the measurement model	P<0.5 indicates significant differences between two paired models.
Incremental fit indices: CFI	> 0.9 suggests a good fit.
Parsimonious fit index: PGFI	>0.5 suggests a good fit.

3.4.4.4 CB-SEM estimation techniques approach

The estimation of a model involves determination of the value of unknown parameters and error terms associated with the estimated values (Weston and Gore Jr., 2006). The possible estimation methods for SEM involve maximum likelihood (ML), least squares (LS), generalized least-squares (GLS), and asymptotic distribution-free (ADF or ADF-WLS) (Weston and Gore Jr, 2006).

In contrast to ML and generalized LS that assume multivariate normality, LS and ADF can be applied for non-normally distributed data analysis. LS does not provide a valid inference to the population from a sample, while ADF can do so if the sample size is large (> 1000) (Bryne, 2013). According to Yuan and Bentler (1998), ADF requires a sample size over 500 even for the simplest model. However, ML is very popular (Hair *et al.*, 2010; Iacobucci, 2009) because it is robust to a moderate violation of the normality assumption of the data set (Anderson and Gerbing, 1988; Hair *et al.*, 2010). ML is defined as “a method of statistical estimation which seeks to identify the population parameters with a maximum likelihood of generating the observed sample distribution” (Lewis-Beck, 1994:153). This study chooses ML estimation due to its flexible approach and can be applied into non-normal distribution data analysis as mentioned above.

Likelihood seeks to find out the probability of a data distribution characterized of parameter θ , given the observation X : $L(\theta|X)$. The likelihood function is a set of proportional probability function: $L(\theta|X) = a * p(x|\theta)$, where a is the proportional constant. In order not to consider the proportional

constant a , the likelihood ratio (LR) is introduced:
$$LR = \frac{L(\theta|X_1)}{L(\theta|X_2)} = \frac{P(x_1|\theta)}{P(x_2|\theta)} \quad (3.26).$$

Denoting S as the observed covariance matrix of the sample, Σ as the implied covariance matrix, their probability density function follow Wishart distribution, named after John Wishart who initially formulated the generalized multidimensional χ^2 (or two-Kchi) distribution in 1928. The likelihood ratio

of S fitting Σ is therefore: $LR = \frac{W(S, \Sigma(\theta), n)}{W(S, S, n)} \quad (3.27)$. Taking the log, the log-likelihood-ratio (LLR) is the

ML estimator. The ML fitting function is obtained by multiplying $-\frac{2}{n}$ and LLR.

3.4.4.5 Mediation Analysis

The causal relationships between antecedents are evaluated by examining the moderation / mediation effects among variables, following the logic proposed by Baron and Kenney (1986).

Mediation effects and indirect effects are often exchangeable (Preacher *et al.*, 2007). Baron and Kenny (1986) state that the mediation effects refers to the causal effects of independent variable X on the dependant variable Y which are transmitted through the mediator M_e , and it's logic can be expressed as (Baron and Kenny, 1986):

$$M_e = a_0 + a_1X + \epsilon \quad (3.28)$$

$$Y = b_0 + c'X + b_1M_e + \epsilon \quad (3.29),$$

where a_0 and b_0 are intercepts, a_1 , c' and b_1 are the coefficients, ϵ represents the regression residuals. Baron and Kenny (1986) propose the causal step strategies by firstly examining the significance of a_1b_1 and secondly c' . The indirect effects occur if the product of a_1b_1 is significantly different from zero. c'

equals to zero indicates that M_e completely mediates the effects of X on Y; otherwise, X regressing to Y is partially mediated by the mediator M_e .

z-tests can be used to determine the significance of the product a_1b_1 . In a case of population estimation, it can be constructed by calculating the standardized errors (Preacher *et al.*, 2007):

$$CI_{1-\alpha}: \hat{a}_1\hat{b}_1 \pm z_{\alpha/2}SE_{\hat{a}_1\hat{b}_1} \quad (3.30),$$

where $SE_{\hat{a}_1\hat{b}_1} = \sqrt{\hat{a}_1^2s_{b_1}^2 + \hat{b}_1^2s_{a_1}^2}$ (3.31). $SE_{\hat{a}_1\hat{b}_1}$ in the equation (3.30) is the Sobel (1982) first-order test, and comparing with the standard normal distribution tests, are represented by the equation (3.28) and (3.29). If zero sets outside of the results by (3.28) (95% confident that the estimations are accurate for α at 0.5), the mediation effects can be considered present.

Moderation effects refer to the causal relationships between X and Y are influenced by a third variable M_o , or the third variable M_o interacts with X regressing to Y (Baron and Kenny, 1986), and it's logic can be expressed as (Preacher *et al.*, 2007):

$$Y = a_0 + a_1X + a_2M_o + a_3XM_o + \epsilon$$

Or (3.32)

$$Y = (a_0 + a_2M_o) + (a_1 + a_3M_o)X + \epsilon$$

where $(a_2 + a_3M_o)$ is the function of the moderator M_o . If a_3 is significantly different from zero, it can suggest that Y regressed on X is significant with the presences of the interactions between the moderator and X (Aiken and West, 1991; Preacher *et al.*, 2007). Similarly, the significance of $(a_2 + a_3M_o)$ can be calculated through (Preacher *et al.*, 2007):

$$SE_{(a_2+a_3M_o)} = \sqrt{s_{\hat{a}_1}^2 + 2s_{\hat{a}_1\hat{a}_3}M_o + s_{\hat{a}_3}^2M_o^2} \quad (3.33)$$

The above discussions show the differences between moderation and mediation approaches by their definitions. A mediator could also be a moderator (Baron and Kenny, 1986). There are two main streams of previous studies on the combination of mediated and moderated effects, where the mediated and moderated effects are presented separately (Donaldson, 2001) or evaluated simultaneously (Edwards and Lambert, 2007; Preacher *et al.*, 2007). Embedded in the theoretical background, the moderated mediation and mediated moderation models can be created when the moderation and mediation effects are examined simultaneously (Edwards and Lambert, 2007; Preacher *et al.*, 2007). The mediated moderation models often require an experimental design and variables are measured in interval/ordinal levels, which are suitable for a small sample size (Preacher *et al.*, 2007). In this study, the moderated mediation effects can be tested, because they can be performed on continuous variables and for a big sample size (Preacher *et al.*, 2007).

a_1b_1 in the equation (3.22) and (3.23) are asymmetrically distributed, and a_1 is independent from b_1 . This suggests that the CIs are not necessary symmetric (Fritz and MacKinnon, 2007; Preacher *et al.*, 2007; Preacher and Hayes, 2008). It is because Sobel's (1982) first order test in the equations (3.31) relies on a normally distributed sample, the percentile bootstrap method that allows the confidential intervals CIs being asymmetric, is commonly used to avoid the Type 1 errors (i.e. incorrectly accepting the null hypothesis of non-indirect effects in this case of mediation analyses) (Fritz and MacKinnon, 2007).

Non parametric bootstrapping refers to randomly chosen samples taken from the initial data set for replacement to the bootstrapping data, repeating this procedure of re-sampling k times (Preacher *et al.*, 2007). This bootstrapping method will calculate the estimations with the replaced sample rather than with the original data set. That is, the new dataset is the distribution of the samples distribution. Preacher and Hayes (2008) argue that k is preferred over than 1000. When the new sample size is large enough, the distribution of a_1b_1 can tend to be normally distributed (Preacher *et al.*, 2007). For this reason, a bootstrap method which re-samples 2000 times is conducted using the software AMOS, with confidence interval set as 0.95.

The values estimated through the percentile bootstrapping method are thereafter sorted from low to high, with the lower and higher boundaries setting between $\alpha/2 \cdot k$ th and $(1 + (1 - \alpha/2) \cdot k)$ th (50th to 1951 th values for repeating 2000 times). The alternative hypothesis of indirect effects cannot be rejected at the 5% level of significance if zero sets outside of the CIs (Preacher *et al.*, 2007).

Percentile-bootstrapping can be further improved with bias-correct percentile bootstrapping method (Preacher *et al.*, 2007) and performed with AMOS. The bias is the difference between the expected testing error and training error (Steck and Jaakkola, 2003). The testing error, also the general error, is the prediction error over an independent sample. The expected testing error is the average of errors that are associated with random, including randomness in the empirical dataset. Training error is the average loss over the training sample (Hastie *et al.*, 2014). It has been found that the training error is smaller than the testing error (Steck and Jaakkola, 2003; Hastie *et al.*, 2014).

Steck and Jaakkola (2003) argue that for a given empirical distribution dataset D , denoting $X = (X_1, X_2, \dots, X_n)$, the empirical data distribution $\widehat{p}(X) = \frac{N(X)}{N}$ (3.34), with $N = \sum_X N(X)$ (3.35). $\hat{\theta}$ is the statistical parameter calculated from D , and its bias is defined through: where $\theta(T)$ is the unknown true estimations of the population, $\theta(p)$ is the associated normalized distribution. It is because $\theta(T)$ is unknown, it is impossible to calculate Bias_T . However, it can be approximated with the bootstrapping bias estimations (Steck and Jaakkola, 2003), where $E(\hat{\theta}(D) - \theta(T))$ is replaced by the average over the bootstrap samples $\langle (\theta(D)^B) \rangle_b$, and p by \hat{p} generated from the empirical data D . $\theta(\hat{p})$ is now called 'plug-in', with its variances calculated through: $\hat{\theta}(\sigma^2) = E(X)^2 - (E(X))^2$ (3.36), and the unbiased one : $\theta_{\sigma^2}^{\text{unbiased}} = \frac{N}{N-1} \hat{\theta}(\sigma^2)$ (3.37). The general bias-correct bootstrapping process is defined therefore as (Steck and Jaakkola, 2003): $\theta^{\text{BC}}(D) = \theta(\hat{p}) - \left(\langle (\theta(D)^B) \rangle_b - \theta(\hat{p}) \right) = 2\theta(\hat{p}) - \langle (\theta(D)^B) \rangle_b$, $b \in (1 \dots k)$ (3.38).

However, $\theta(\hat{p})$ in most cases it is non-linear, thus the biases do not disappear automatically (Steck and Jaakkola, 2003). Bias-corrected maximum likelihood estimations can be used to address this problem.

The principle maximum of entropy for the discrete data (empirical data) is defined as (Steck and Jaakkola, 2003; Conrad, 2013): $H(P(X)) = -\sum_x p(x)\log(p(x))$ (3.39), with $p(x)$ representing the true (unknown) probability density function.

Thus a higher $H(P(X))$ suggests less information over a distribution or more uncertain situation (Conrad, 2013). The goal is to learn the population with samples using a Bayesian network, i.e. give the empirical data to calculate the probability of the true parameters of a population. The test error of the learned model is defined (Steck and Jaakkola, 2003): where T is Taylor approximation, m denoting the known Bayesian network structure, $\hat{p}(x|m)$ represents the conditional probability of presence of X by giving m . $p(x)$ is unknown, but is approximated to the training data: $T(\hat{p}, m) = -\sum_x \hat{p}(x|m)\log(\hat{p}(x|m))$. The bias is: $T\text{Bias} = -H(p(X|m)) - (\frac{1}{2N}|m|-1) + O(\frac{1}{N^2}) - (-H(p(X|m)) - (\frac{1}{2N}|m|-1) + O(\frac{1}{N^2})) \approx \frac{1}{N}|m|$ (3.40) (Steck and Jaakkola, 2003).

(e.g. the differences between Bootstrapping and empirical data are: $H(P(x)) = \langle H(D^B) \rangle - H(\hat{p})$, where

$$\langle H(D^B) \rangle = \frac{1}{B} \sum_{b=1}^B H(D^b(X)) = -\sum_x \langle \frac{\text{Var}(X)}{N} \rangle \log \frac{\text{Var}(X)}{N};$$

\therefore E.g. $\text{Var}(X) \sim \text{Binominal}(N, \hat{p}(X))$;

\therefore Its second order Taylor expression: $\sum_x L(\frac{\text{Var}(X)}{N}) = \frac{\text{Var}(X)}{N} \log(\text{Var}(X)) = H(\hat{p}(x)) - \frac{1}{2} \sum_x L''(\hat{p}(x)) (\frac{\text{Var}(X)}{N} - \hat{p}(X))^2 + O(\frac{1}{N^2}) = H(\hat{p}(x)) - (\frac{1}{2N}|X|-1) + O(\frac{1}{N^2})$

The analysis follows the logic of above discussions, and additionally uses structural equation modelling with latent variables (Muller *et al.*, 2005; Hopwood, 2007). An advantage of incorporating latent variables in contrast to observed variables is that it can ameliorate the reliability and method effects on the mediation and moderation models (Hopwood, 2007). The measurement errors associated with one particular observed variable that is specified loading toward the latent variable is unlikely to be shared

with other observed variables that are equally loading toward the same latent variable, since latent variables seek to measure the overall desired effects (Hopwood, 2007).

The results generated from study one indicate that both online trust and perceived critical mass are important antecedents to intention to contribute knowledge within online communities. In order to provide insight on what the latent factors online trust and the perceived critical mass really are, the following study two takes an inductive approach, and study three precedes an abductive approach. In other words, study two and three are independent from study one, but can provide richer information on results generated from the study one. The three studies seek to investigate the antecedents of intentional contribution behaviours online, but with different worldviews. This is in agreement with the nature of mixed-methods (Venkatesh *et al.*, 2013). The following sections will further discuss the research methods for studies two and three.

3.5 Methodology of study two - evaluating online trust

3.5.1 Research aims

Study one sought to understand antecedents of intention to contribute knowledge online, embedded in the theory of DTPB. Results identified trust in online forums as an antecedent, together with perceived critical mass positively influence the intention to contribute knowledge online. It is recognized that the processes by which trust is developed and undermined is not revealed by study one.

Study two seeks to contribute to the understanding of the antecedents and implications for marketing management of the development of institutional trust, and in particular to learn more about how the components of trust change over time. Given that consumers frequently turn to online forums to assess the trustworthiness of a potential supplier; it may seem surprising that relatively little is known about how the components of institutional trust change over time. This study aims to learn more about the processes of trust formation within the context of peer-to-peer electronic media.

The study combines inductive and deductive aims. The inductive aim is to build theory that is then tested using deductive, quantitative approaches. The research problem is specifically contextualized to a fast growth technology based brand operating in an environment where peer-to-peer comments can rapidly contribute to, or undermine trust.

Although the aim is to develop and test a generalizable theory, it is recognized that the processes by which trust is developed and undermined is context- and situation-specific, therefore there would be limitations in generalizing beyond the sector studied, or beyond the context of the communications technologies available at the time.

3.5.2 Data collection

Data is collected for the period 2005-2010. This time series approach does not allow comments of unique individuals to be tracked, so the study cannot be described as longitudinal. The time series approach allows aggregate comments to be compared between years. The research framework does not explicitly incorporate assessments of trust that might be pervasive in society, rather than specific to the focal organization being studied. It is also possible that changes in trust over time may reflect differences in the composition of reviewers at each period.

A brand is sought that has grown virally through social network media, and has the capacity to be undermined through those same media. In order to explore the dynamics of trust, the brand should at first sight appear to have gone through some type of cycle of trust over a period of 5-10 years. The focal organization should be one that is widely known and for which large volumes of historically archived comments are available.

After considerable review and discussion with experts in the field, Skype is chosen as the focal brand. Skype provides voice over internet protocol (VoIP) telephony service. It is founded in 2003 and experienced rapid growth, based largely on word-of-mouth recommendation through social network media. According to Soomro and Asfandiyar (2010), the global VoIP industry went through a rapid

development from 2005 to 2010, particularly during a period of global recession from 2008-2010. However, as Skype has grown and been acquired by larger corporations - eBay, and subsequently Microsoft, there is a suggestion that trust in the brand may have been undermined by association with its new owners (Halliday, 2011).

A review of news media indicates that during the period of study, a number of events had occurred which might have influenced users' evaluation of trust in Skype. These include widely reported major system failures (Arak, 2007 http://heartbeat.skype.com/2007/08/what_happened_on_august_16.html), litigation over the company's technology (in which Skype has at various times been a plaintiff and a defendant), and changes in ownership which may have transformed perceptions of the company in many people's minds from a group of young, innovative, hardworking and challenging entrepreneurs to an anonymous, profit seeking multinational corporation (Halliday, 2011).

The analysis is based on comments provided by 234 reviewers on 3 peer review websites (www.dooyoo.com; www.ciao.com; www.reviewcentre.com). These peer-to-peer sites are independently moderated and not tied to a particular company being reviewed. The owners of the sites generate income from adverting and "click-through" payments. The researcher examines all reviews relating to Skype posted on the three websites between 2005 and 2010. Each review describes a personal experience, many of which make direct or indirect references to trust in the brand. The review sites do not provide demographic or behavioural data in a structured manner, and there has been concern about the authenticity of reviews left at open access review sites, with claims that reviews may include critical comments submitted by competitors, and complementary reviews submitted by the management of the reported business (Garcia, 2007).

It is difficult to establish the representative validity of the reviewers used in this study. However, it has also been noted that there is a tendency for strongly negative reviews to be roughly balanced by strongly positive ones (Grandcolas *et al.*, 2003). Furthermore, the inductive approach sought depth of informed comment rather than an ability to make inferences to the population of Skype users as a whole.

3.5.3 Method

This study combines phenomenological and quantitative approaches. The phenomenological underpinning seeks to understand – through systematic analysis – individuals' structure of consciousness through their own personal accounts rather than through theoretical structures imposed by the researcher (Sokolowski, 2000).

The study initially sought to build a theory of trust on the basis of comments left by users of publicly accessible online peer-review sites. Observation of social network media with a view to obtaining meaning has led to the development of a range of techniques, such as transaction log analysis (Jansen *et al.*, 2000), verbal protocol analysis (Nahl and Tenopir, 1996), and “webnography” or “virtual ethnography” (Morton, 2001). A spectrum of webnographic approaches exist from purely observational to full participatory (Morton 2001:6). The approach adopted here is distanced, implying evaluation of texts and observation of social interaction in online environments, but not participation in them.

It has been suggested that webnographic approaches are particularly suitable for conducting research about hi-tech products among “leading edge” and technology-savvy consumers (Puri, 2007). Furthermore, an anonymous internet environment can encourage contribution of impulsive comments without contributors fearing reprisals, creating a source of data that may be more expressive and potentially more credible than if comments were moderated by the need to conform to subjective norms and the prospect of retribution for breaching those norms (Puri, 2007). Against this, it has been suggested that the presence of dishonest and intentionally malicious comments in online fora has led many users to trust online word-of-mouth less than they would trust face-to-face word-of-mouth (Keller and Berry 2006).

The difficulty of obtaining representative samples using webnographic techniques have been noted (Kozinets, 2006). Against this, it can be objected that qualitative research is essentially about discovering meanings attributed by people who are engaged in the topic of the study, and what counts is depth of understanding rather than representativeness of those being studied. The results of the inductive phase of

this study can only allow a generalization to theoretical propositions rather than to a statistically reliable prediction of behaviour (Yin, 1984).

The outcome of the initial inductive phase of the research informed a series of deductive tests of an emergent theory. This entailed using scores for data collected in the inductive phase and examining associations within the data using tests of statistical significance. In particular, any suggestion in the inductive phase about cyclical patterns of the components of trust was explored further and subjected to tests to verify their significance.

Four items emerge from the inductive coding namely attitude, benevolence, integrity and predictive. The descriptive data analysis shows that the terms emerging from the qualitative analysis are different from the normal distribution and multicollinearity across the data set that is the important assumption for the linear regression model. However, it is noted that predictors of a dependant variable are often highly correlated and/ or collinear in the social sciences (Hair *et al.*, 2010).

Principle component analysis (PCA) (see the section 3.4.4.2) seeks to reduce the abundant information in the predictors that can be the cause of multicollinearity (Hair *et al.*, 2010). To predict the dimensionality of online institutional trust, the commonly applied model in machine learning namely support vector machine (SVM) for classifications is thereafter conducted. Different to the discriminant analysis that requires the multivariate normal, SVM can map the nonlinear relationships (introducing a kernel function) between the predictors and response variable by projecting the multiple dimensional patterns into the two half spaces (Caragea *et al.*, 2001). For a non-separable one class learning problem, the slack weights are introduced so that the hyperplanes that separate the predictors are maximized but with a penalty cost function associated. SVM is therefore able to address the issue of multicollinearity because the relationships between the variables are studied in the high dimensional spaces.

The evolution of online institutional trust is analysed with the recurrent neural network for time series. Recurrent neural network is useful when the size of the input variables is unequal (Nielsen, 2015), which

hence can deal with the measurement problems. It is this case where the numbers of coding for the predictors vary in year.

3.5.4 Data analysis

3.5.4.1 Coding

The form of analysis used in the initial inductive phase is based on the principles proposed by Miles and Hubermann (1994), Strauss and Corbin (1998), and is iterative in nature. Data collection and analysis are consciously combined, and initial data analysis is used to guide ongoing data collection and coding. Reviews are coded and analysed using nVivo software. This allows the researcher to identify associations between themes of comments, and to supplement this with contextual data introduced by the researcher.

To develop initial themes, free nodes are coded using themes derived from the first 10 randomly selected reviews. Free nodes that shared common underlying ideas were merged into tree nodes. The list of tree nodes gradually expands as more reviews are analysed and represented the list of categories and themes that emerged from the coding. Following the recommendation of Gibbs, coded nodes are arranged into a node tree. Tree nodes are located at the top, representing emerged themes, beneath, which are lists of child nodes and sub-tree nodes (Gibbs, 2002).

Each node is scored on a scale appropriate to that node. For example, comments relating to the emerged theme of ability are rated on a scale from 1 (“I do not believe they can fix my problem”) to 5 (very positive comments, including statements such as “brilliant”, “fantastic”, “completely” cited in respect of the abilities of the company). It should be recognized that the node hierarchy structure is subject to the researcher's subjective interpretations, although inter-rater reliability was measured (discussed below).

The following tabulation 12 and figure 9 illustrates the scales used for rating the node “ability”:

Table 12: An example of scales used for rating the node “ability”

Rating	Meaning	Illustration
1	Very weak	“I do not believe they can fix my problem”
2	Weak	“They could have done better”
3	Average	“Not sure that they fixed the problem”
4	High	”They fixed the problem quickly”
5	Very high	Statements such as “brilliant”, “fantastic”, “completely” cited in respect of the abilities of the company

Analysis is undertaken for those comments, which have relevance to the specified dimension of trust. Therefore, if a posted comment contains no relevant connection to that dimension, it is not scored. Comments are scored independently by two researchers. Based on a sample of 50 posted comments, there is an inter-rater reliability score of .877. To give an indication of the numerical coding applied to each comment, the actual score given to each comment is shown for the example verbatim comments given in the table 13.

Four themes emerge from the process of initial free coding and subsequent development of tree nodes and comprised: ability, benevolence, integrity and predictability. In addition, the dimension of “overall trust” is identified. An example of year 2009 shows in the figure 9. The four themes are illustrated in Appendix 2 with examples of verbatim comments.

Figure 9: An example of year 2009: ability scaling

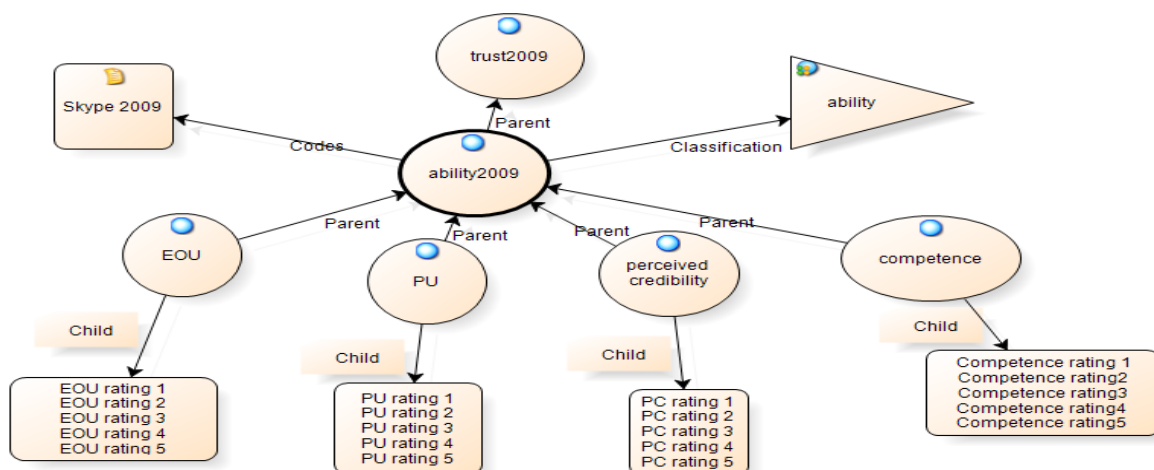
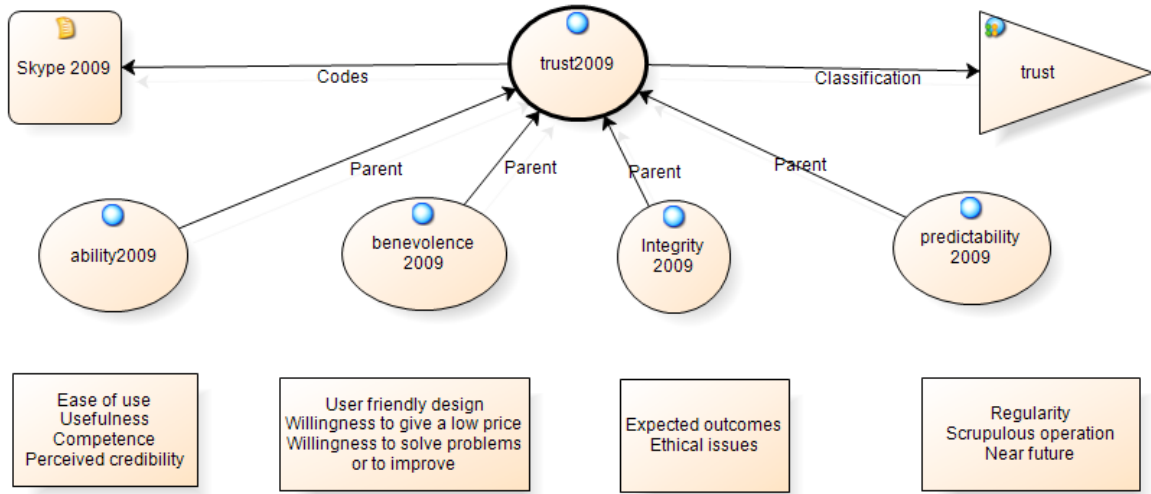


Figure 10: an example of year 2009: emerging themes



3.5.4.2 SVM

The basic idea of support vector machine (SVM) is to use some fixed nonlinear function (Kernel function) to map the high dimensional input X so that the linear relationship between the target y and the input X is able to be defined:

$f(x, \omega) = \langle w, x \rangle + b$, with $w \in X, b \in \mathfrak{R}$ (3.41), \langle, \rangle is the dot production within the input; b is the bias term .

The quality of the above estimations is measured in this study with the ε - intensive loss function (Cortes and Vapnik ,1995): $\varepsilon = 0, \text{if } |\xi| \leq \varepsilon; = |\xi| - \varepsilon, \text{otherwise}$. The loss function of (3.41) is therefore defined as:

$$L_g(y, f(x, \omega)) = \begin{cases} 0, & \text{if } |y - f(x, \omega)| \leq \varepsilon \\ |y - f(x, \omega)| - \varepsilon, & \text{otherwise} \end{cases} \quad (3.42)$$

The objective is to minimize $\frac{1}{2} \|\beta\|^2 + C \sum_{i=1}^m (\xi_i + \xi_i^*)$ (3.43), subject to $\begin{cases} y_i - \langle \omega, x_i \rangle - b \leq \varepsilon + \xi_i \\ \langle \omega, x_i \rangle + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \\ C > 0 \end{cases}$, where

C is a constant deviation trade-off between the optimised flat and calculated f ;

(3.43) is submitted into the Lagrange function (Smola and Schölkopf, 1998) for the nonlinear quadratic model:

$$L(a, \beta, \omega) = \frac{1}{2} \|\beta\|^2 + C \sum_{i=1}^m (\xi_i + \xi_i^*) - \sum_{i=1}^m a_i (\varepsilon + \xi_i - y_i + \langle \omega, x_i \rangle + b) - \sum_{i=1}^m a_i^* (\varepsilon + \xi_i^* - y_i - \langle \omega, x_i \rangle - b) - \sum_{i=1}^m (\eta_i \xi_i - \eta_i^* \xi_i^*) \quad (3.44),$$

and solving for the optimization of dual problem by vanishing the primal factors:

$$\begin{aligned} \partial_b L(a, \beta, \omega) &= \sum_{i=1}^m (a_i^* - a_i) = 0; \\ \partial_\omega L(a, \beta, \omega) &= \omega - \sum_{i=1}^m (a_i - a_i^*) x_i = 0; \\ \partial_{\xi_i^{(*)}} L(a, \beta, \omega) &= C - a_i^{(*)} - \eta_i^{(*)} = 0; \end{aligned} \quad (3.45)$$

Submitting the three equations in (3.45) into (3.44), it yields to maximize

$$L(a, \beta, \omega) = \frac{1}{2} \sum_{i,j}^m (a_i - a_i^*)(a_j - a_j^*) \langle x_i, x_j \rangle - \varepsilon \sum_{i=1}^m (a_i + a_i^*) + \sum_{i=1}^m y_i (a_i + a_i^*) = 0$$

Or to maximize:

$$\begin{aligned} & -\frac{1}{2} \sum_{i,j}^m (a_i - a_i^*)(a_j - a_j^*) \langle x_i, x_j \rangle \\ & - \varepsilon \sum_{i=1}^m (a_i + a_i^*) + \sum_{i=1}^m y_i (a_i + a_i^*) \end{aligned} \quad (3.46), \text{ subject to } \sum_{i=1}^m (a_i^* - a_i) = 0, \quad a_i^*, a_i \in [0, C].$$

The optimised solutions to (3.46) produce the support vectors for the decision boundary that minimized maximize or maximized minimize distance between X. More generally, (3.46) can be expressed with the

$$f(x) = \sum_{i=1}^{n_{sv}} (a_i - a_i^*) K(x_i, x) \quad (3.47), \text{ where } n_{sv} \text{ is}$$

Kernel function that transfer the inner production within X:

the numbers of support vectors, subject to $a_i^*, a_i \in [0, C]$.

Following the suggestions from Ng (2016), the k-fold cross validation method is preferred with k=10 in order to make valid results from SVM:

- Randomly split the dataset into 10 folders S_1, \dots, S_{10}
- For each model i :

For $j = 1, \dots, 10$, testing the models on the sample size $S_1 \cup \dots \cup S_{j-1} \cup S_{j+1} \cup \dots \cup S_{10}$, excepting for S_j . This results hypotheses h_{ij} that are further tested on S_j . The estimated general error of

each model is calculated:

$$\hat{\epsilon}_{ij} = \frac{\sum_{i,j} (\hat{y}_{ij} - y_{ij})^2}{10} \quad (3.48).$$

- The smallest $\hat{\epsilon}_{ij}$ suggests that the raw dataset is best fitted with the Mdl_i , that is finally tested with the whole sample S .

Four kernel functions are accessible: linear dot product $K(x_i, x_j) = x_i \cdot x_j$; Gaussian (RBF) $\exp(-\gamma |x_i - x_j|^2)$; inhomogeneous polynomial $(\gamma x_i \cdot x_j + C)^p$, $p = 2, \dots$; sigmoid $\tanh(\gamma x_i \cdot x_j + C)$. One important advantage of using SVM is that the nonlinear relationships mapping X can be transferred with the Kernel function and solved within the optimal dual quadratic problem.

3.5.4.3 Using recurrent neural network in time series

Recurrent neural network for time series is recognised as an unsupervised method for mapping the complex nonlinear relationships between the inputs, the institutional trust scores by year (2005-2010). The objective is to predict the overall trust score in the coming years. The N samples are divided into three datasets: 70% for the training ($0.7N$), 15% for validations ($V=0.15N$) and 15% for testing ($T=0.15N$). Validations are used to measure the generalization of the results from the training where the network is adjusted with the back propagation algorithms. Testing is independent from the training and measures the performance of the network during and after training.

The model in this study is constructed as:

$y(t) = f(y_{t-1}, \dots, y_{t-d})$ (3.49), where d is the time delays by year (=6);

$$\hat{y} = \sum_{l=1}^6 w_j z^l + b \quad (3.50)$$

$$z^l = \sum_{j=1}^6 \sigma(w_{jk} y_{t-d} + b_j) \quad (3.51),$$

where w_{jk} represents weights from the neuron k in the previous layer to the neuron j in the next layer; b is the bias and σ is an nonlinear function.

For every input vector presented in the training sample, the output is compared with the desired vector.

The goal is hence to minimize the learning errors. The cost function (or error) of the model is defined

using the gradient batch: $E = \frac{1}{2} \sum_{N-V-T+1}^N (\hat{y} - y_{train})^2$ (3.52). The performance of the model is to compute the

mean squared error (MSE) with the lower value indicating less errors: $\frac{E}{\|S\|}$, where the set S representing the size by training, validation or testing. The output is evaluated with the regression adjusted R^2 value, for which a value closes to 1 referring a correlation between years. The computation of the adjusted R

squared is:

$$R - square = 1 - \frac{\sum_{i=1}^{N-1} w_i (\hat{y}_i - y_i)^2}{\sum_{i=1}^{N-6} w_i (\hat{y}_i - \bar{y})^2} \quad (3.53).$$

The σ function that applied on every neuron in the hidden layers is sigmoid function in this study:

$\sigma(z_j^l) = \frac{1}{1 + e^{-z_j^l}}$ (3.54), which leads to $\begin{cases} = 1, \text{if } \sigma(z_j^l) \geq 0.5 \\ = 0, \text{otherwise} \end{cases}$. Sigmoid function is also known as the logistic function, which can class the neuron by ‘weak=0’ and ‘strong=1’.

The model is embedded in the ground truth value of y , and can be performed with the backpropagation algorithm, cited with the proposition by Nielson (2015). The following figure 11 explains the neural network model in this study (see figure 11).

The estimated \hat{y} by the neural network is compared with the ground truth value of y , and adjusts the local weight and bias so that such differences are minimized. There are cost functions with respect to the

changes in w ($\frac{\partial C}{\partial w_{jk}^l}$) and b ($\frac{\partial C}{\partial b_j^l}$). The cost function computed in the output layer with respect to an

individual neuron is understood: $C = \sigma(z_j^l + \Delta z_j^l)$. This suggests how the neuron j in the layer l interprets

the information received to adjust w and b , and report $\sigma(z_j^l + \Delta z_j^l)$ instead of $\sigma(z_j^l)$. Δz_j^l should be

optimised to be small. The individual neuron error is defined : $\delta_j^l = \frac{\partial C}{\partial z_j^l}$ (3.55) (e.g. $\frac{\partial C}{\partial z_j^l} = \frac{\partial C}{\partial z_j^l} \Delta z_j^l$

$= \delta_j^l \Delta z_j^l$). The individual neuron error therefore measures the ability of an individual neuron in handling

Δz_j^l . If $\delta_j^l = \frac{\partial C}{\partial z_j^l}$ is large either positive or negative, this neuron reports a small Δz_j^l . Otherwise, the

individual neuron j in the layer l is not able to improve Δz_j^l .

Remind that $\sigma(z_j^l)$ is associated with the neurons activations. The neuron error term in the output layer L

is : $\delta_j^L = \frac{\partial C}{\partial a_j^L} \sigma'(z_j^L)$ (e.g. $\delta_j^L = \frac{\partial C}{\partial a_j^L} \frac{\partial a_j^L}{\partial z_j^L}$) (3.56). If $\frac{\partial C}{\partial a_j^L}$ is small, the cost changes are not heavily

dependent on the activations. The computations depend on the cost function. For example, the gradient

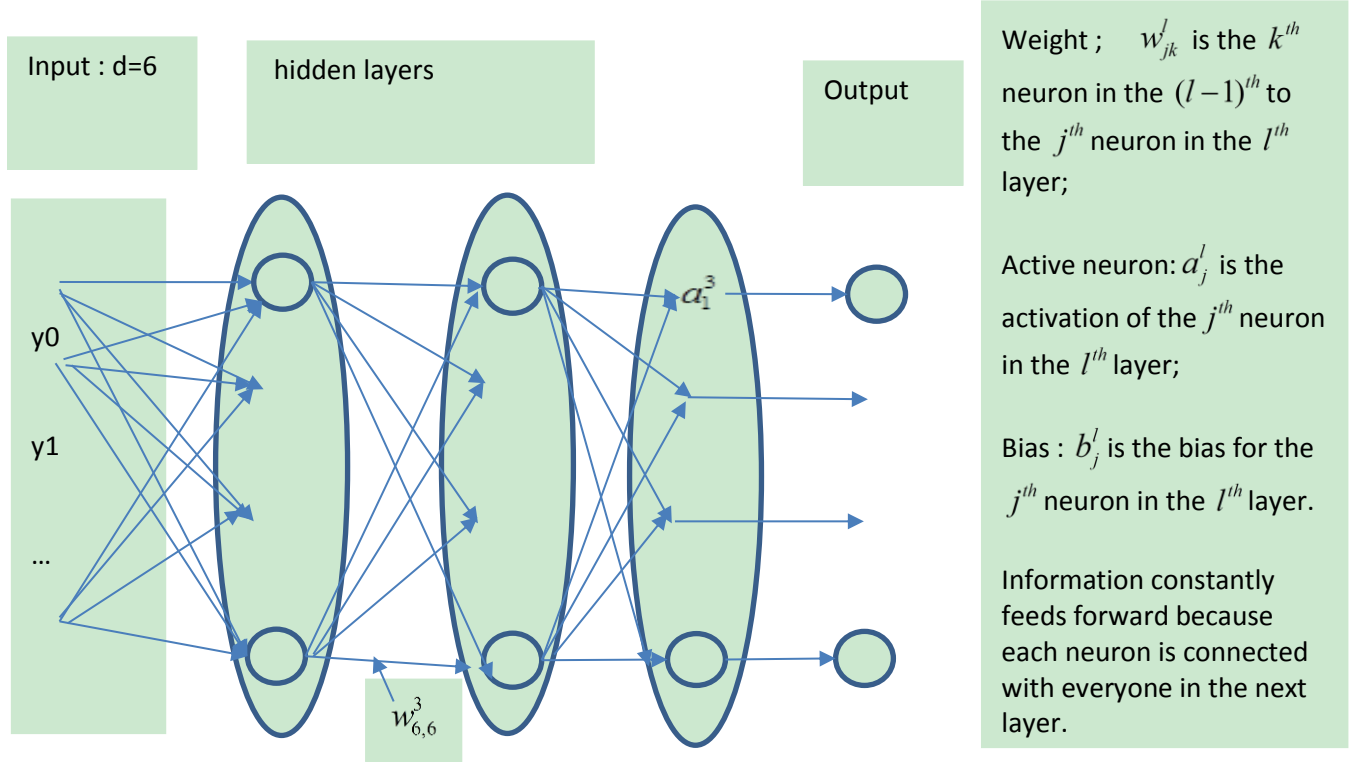
method to define the cost function in term of activations: $C = \frac{1}{2} \sum_j (y_j - a_j)^2$; $\frac{\partial C}{\partial a_j^L} = a_j - y_j$;

$$\frac{\partial a_j^L}{\partial z_j^L} = \frac{1}{\frac{\partial z_j^L}{\partial a_j^L}} = \frac{1}{\sum_k w_{jk}^L}$$

. In other words, the error term of the neuron j in the layer L is able to be calculated

with output activations and the input weight.

Figure 11: Recurrent neural networks for time series



$$a_j^l = \sigma(\sum_k w_{jk}^l a_k^{l-1} + b_j^l) ; z_j^l \equiv \sum_k w_{jk}^l a_k^{l-1} + b_j^l$$

In the situation in which the error term in the next layer is known, the neural network can backward

$$\begin{aligned} \delta_j^l &= \frac{\partial C}{\partial z_j^l} = \sum_k \frac{\partial C}{\partial z_k^{l+1}} \frac{\partial z_k^{l+1}}{\partial z_j^l} = \sum_k \frac{\partial z_k^{l+1}}{\partial z_j^l} \delta_k^{l+1} \\ &= \sum_k \frac{\partial(\sum_j (w_{kj}^{l+1} \sigma(z_j^l) + b_k^{l+1}))}{\partial z_j^l} \delta_k^{l+1} \\ &= \sum_{k,j} w_{kj}^{l+1} \delta_k^{l+1} \sigma'(z_j^l) \end{aligned}$$

compute the error term in the current output layer: , or defined in

the matrix form $\delta^l = ((w^{l+1})^T \delta^{l+1}) \odot \sigma'(z^l)$ (3.57) . (3.57) is therefore able to measure the error term in

the layer (l-1), giving the error term in the layer l. Together with (3.56), the error term in each layer

within the neural network that is associated with the input weighted activations can be computed.

Having computing δ_j^l , the partial derivation of $(\frac{\partial C}{\partial w_{jk}^l})$ and $(\frac{\partial C}{\partial b_j^l})$ is measured in term of δ_j^l .

$$\frac{\partial C}{\partial w_{jk}^l} = \frac{\partial C}{\partial z_j^l} \frac{\partial z_j^l}{\partial w_{jk}^l} = \delta_j^l \frac{\partial(a_j^l w_{jk}^l + b_j^l)}{\partial w_{jk}^l} = \delta_j^l \sigma'(a_j^l) = \delta_j^l a_k^{l-1} \quad (3.58)$$

$$\frac{\partial C}{\partial b_j^l} = \frac{\partial C}{\partial z_j^l} \frac{\partial z_j^l}{\partial b_j^l} = \delta_j^l \frac{\partial(a_j^l w_{jk}^l + b_j^l)}{\partial b_j^l} = \delta_j^l \quad (3.59)$$

(3.58) and (3.59) suggest that the cost function can be improved by adjusting the individual neuron term and the active neuron in the previous layer that is associated with the weight input to that neuron. Because the σ function is applied on each neuron in the hidden layer, a smaller value of w_{jk}^l will lead to a more rapid decrease in the cost function. A learning rate $n=0.1$ is given to the difference of weight ($\Delta w^l = -n\delta^l a^{l-1}$) so that the decrease in the cost function is slower.

As explained above, perceived critical mass, in addition to trust, is another determinant of intention to contribute online. The following sections seek to provide a richer understanding of perceived critical mass in the context of online contributions.

3.6 Methodology of study three - the role of critical mass in sustaining online forums

3.6.1 Research aims

The results generated from study one indicate that perceived critical mass has a positive influence on subjective norms and intention to contribute online. The theory of critical mass (Oliver and Marwell, 1988) highlights that small groups of members can evoke mass collective actions, suggesting a phase transition emerges after a critical point. Critical mass members are initial contributors who pay the set up cost of public goods, and who in general have more resource to contribute (Oliver and Marwell, 1988).

Results from study one embedded in the online survey with 910 responses support two essential principles proposed by the theory of perceived critical mass: 1) the size of network is often associated

with the possibility of the emergence of hubs-the critical mass members; 2) the critical mass members often play the role of “bridge” between the communications of members.

Results highlight the importance of critical mass members who can further reveal the structure of networks that is evolving over time. To gain a richer understanding on the influences of critical mass members on the structural dynamics of online networks, it is explored with an analysis and virtualization of network structures within which the role of critical mass members is examined.

3.6.2 Data collection

Wasco *et al.* (2009) examine single inter-organizational networks for studying social structure in an electronic network of practice. In the studies of Westland (2010), a picture was taken to show the structure of a social network (Facebook). However, online forums are dynamic with continual addition and attrition of members, which requires a dynamic view of data mining from online forums.

Network analyses have been undertaken on the open access “stack overflow” data provided by SNAP, a database hosted by Stanford University. “Stack Overflow” (stackoverflow.com) is an interest-oriented online discussion forum where members can share knowledge on programmes such as using JAVA and Python languages. From July 2008 to May 2014, a total of 19,881,020 posts are published with a mean of around 336,966 posts per month.

Stack overflow provides a data dump of user generated data such as the list of questions and answers, which are available online (see web link below). The total numbers of questions are 7,214,697 within which the top questions asked are for “Java” programming languages (632, 493). Given that the compressed data dump by stack overflow is an extremely big file (5.2GB compressed), the study has therefore taken the data dump of questions and answers related to the Java category, which is publicly available by SNAP. Having matched with the answer lists associated with the ‘Java’ question list, the final data set contains 363,147 lines. That is, questions without answers are not included for study.

To construct a ‘contribution’ network, it is firstly necessary to identify whether edges are present between nodes. In this study, each ID name represents one node/actor, and one edge is considered between the poster and the responder (Wang and Dai, 2009). In other words, no matter the frequency of exchanged messages which may depict the density of exchanging activities between two nodes, there is exactly one edge between them. It is noted that different ID names represent different nodes. The situation is neglected when one node creates more than one ID name. By doing this, the un-directed and un-weighted network is constructed over time.

The data downloaded from SNAP represents a direct network, with edges toward and out-of each node. The numbers of edges toward a node is called in-degree, and the numbers of edges leaving a node is named out-degree. The sum of the in-degree and the out-degree of a node suggest its strength of connectivity within the network (Diestel., 2005). The directed network is interpreted as un-direct one in the visualisation software as the edge direction is ignored. The main interest of doing so is that the objective of this study seeks to understand the role of critical mass members (different nodes type) within a network, rather than to explore the strength of interactions between nodes, which is often studied in a directed and weighted network.

To date, research on complex networks is limited to understanding some specific characteristics of particular types of network such as the power law tails in networks of information transition, and empirical data is often collected for a particular purpose because complex network theory is still in its infancy (He *et al.*, 2008). According to Watts (2004), the major challenges of research in the new science of networks leads to the interpretations of findings generated from the analysis of different types of networks.

3.6.3 Method

The structure of a successful online forum in terms of the connectivity associated to nodes (i.e. degree distribution). It is argued that a mass phenomenon is more likely to be studied within a successful network. The study seeks to provide the possible answer to the implementation of theory of critical mass

within online forums by taking a dynamic view. Hence, it firstly examines the structure of the large- scale size of online forum under study, and understands whether it is relatively safe by proposing that it is a successful online forum.

The maximum likelihood (ML) analysis proposed by Clauset *et al.* (2009) is the main technique to fit the degree distribution, which is the main method to explore the structure of a network (e.g. Newman, 2005). Hubs/critical mass members are identified in terms of their connectivity, and their roles in the collision of membership are estimated through *attending* and *random* attacks. In contrast to a random attack that randomly removes members from the network; an attending attack on the network deletes hubs in an increasing fraction (Albert *et al.*, 2000). Simulations have sought to understand under what conditions the network under study is no longer connected, so that the network is broken down and its functionality of self-sustaining is no longer ensured. By comparing the conditions for attending and random attacking, the roles of critical mass members are demonstrated.

A scale-free network continuously grows and develops into a stable state (Barabasi and Albert 1999). Having examined the general structure of the online forums being seen as a network, study three further seeks to investigate the mechanisms that govern the evolution of the network. The critical point is calculated embedded in the method proposed by Cohen *et al.* (2002). This allows analysing four successive stages, represented by far before, near, above and far above the critical point. Hubs (top 50 critical mass members in terms of their numbers of connections) are identified in each stage, in order to investigate their roles in the evolution of online forums.

3.6.3.1 Exploring the structure of online forums

The structure of a network can be examined by the degree distribution (e.g. Clauset *et al.*, 2009). Research conducted by Albert and Barabasi (1999) shows that the degree distribution of scale-free network is characterised by the power law, $P(x) \propto x^{-\gamma}$, $x > 0$, where γ is the scaling or exponent parameter; and x represents the degree (connections) associated to a node. In a scale-free network, a

majority of connections within the network is held by a small part of members, while a majority of members have few or no connections (Albert and Barabasi, 1999).

However, few empirical data follow the power law for all degree (x) value, and more often only degrees greater than the minimum degree ($x > x_{\min}$) obey the power law distribution (Newman *et al.*, 2006; Clauset *et al.*, 2009). Maximum likelihood (ML) is argued as a more accurate method to estimate the scaling parameter in a power law distribution (Wasserman, 2003; Clauset *et al.*, 2009), which is therefore adopted in this study.

Clauset *et al.* (2009) considered separately for cases where x is continuous or discrete, and propose different maximum likelihood estimators (MLEs) of the exponent parameter γ . MLEs for the continuous

x is: $\hat{\gamma} = 1 + n \left[\sum_{i=1}^n \ln \frac{x_i}{x_{\min}} \right]^{-1}$ (3.60), where $x_i, i=1,2,3,\dots,n$, which are the observed degree values satisfying $x_i \geq x_{\min}$.

(E.g. derived from Hill estimator, the probability density function of the power law distribution can be written as: $p(x) = \frac{\gamma-1}{x_{\min}} \left(\frac{x}{x_{\min}} \right)^{-\gamma}, x > x_{\min}$).

$$\begin{aligned} \text{MLES(L)} &= \sum_{i=1}^n \left[\ln(\gamma-1) - \ln x_{\min} - \gamma \ln \frac{x_i}{x_{\min}} \right] \\ &= \sum_{i=1}^n \left[\ln(\gamma-1) - \ln x_{\min} - \gamma \ln \frac{x_i}{x_{\min}} \right] \\ &= n \ln(\gamma-1) - n \ln x_{\min} - \gamma \sum_{i=1}^n \ln \frac{x_i}{x_{\min}} \\ &\because \partial_{L/\partial \gamma} = 0 \end{aligned}$$

$$\hat{\gamma} = 1 + n \left[\sum_{i=1}^n \ln \frac{x_i}{x_{\min}} \right]^{-1}, \text{ where } x_i, i=1,2,3,\dots,n$$

$$\sigma = \frac{\hat{\gamma} - 1}{\sqrt{n}} + O\left(\frac{1}{n}\right),$$

where the big O notation denotes the sample size bias of order which is smaller than $1/n$, and can be neglected when $n > 100$ as it is minor.)

MLEs for the discrete case is: $\hat{\gamma} \cong 1 + n \left[\sum_{i=1}^n \ln \frac{x_i}{x_{\min}^{-1/2}} \right]^{-1}$ (3.61). It is because γ follows the normal

distribution, the standard error is: $\sigma = \frac{1}{\sqrt{n \left[\frac{\zeta''(\hat{\gamma}, x_{\min})}{\zeta(\hat{\gamma}, x_{\min})} - \left(\frac{\zeta'(\hat{\gamma}, x_{\min})}{\zeta(\hat{\gamma}, x_{\min})} \right)^2 \right]}}$ (3.62), where $\zeta(\hat{\gamma}, x_{\min})$ represents

Hurwitz zeta function. For both cases, the statistical error on $\hat{\gamma} \rightarrow 0$, when n is over 100.

(E.g. if $f(x)$ is differentiable, it's indefinite integral $F(x)$ satisfies $F'(x) = f(x)$.

$$\int_{x_{\min}^{-1/2}}^{\infty} f(t) dt = \sum_{x=x_{\min}}^{\infty} f(x) + \frac{1}{24} \sum_{x=x_{\min}}^{\infty} f''(x) + \dots$$

$$= \sum_{x=x_{\min}}^{\infty} x^{-\gamma} + \frac{\gamma(\gamma+1)}{24} \sum_{x=x_{\min}}^{\infty} x^{-\gamma-2} + \dots$$

$$= \zeta(\gamma, x_{\min}) \left[1 + O(x_{\min}^{-2}) \right]$$

$$\because x^{-2} \leq x_{\min}^{-2}$$

$$\zeta(\gamma, x_{\min}) = \frac{(x_{\min} - \frac{1}{2})^{-\gamma+1}}{\gamma+1} \left[1 + O(x_{\min}^{-2}) \right]$$

$O(x_{\min}^{-2})$ is negligible,

$$\hat{\gamma} \cong 1 + n \left[\sum_{i=1}^n \ln \frac{x_i}{x_{\min}^{-1/2}} \right]^{-1}$$

Minimum degree

The above discussions indicate that the lower boundary of x_i , x_{\min} , should be firstly decided in order to estimate the accurate scaling parameter $\hat{\gamma}$ (Clauset *et al.*, 2009). Choosing a value for x_{\min} is important. The estimated $\hat{\gamma}$ will be seriously deviated from the true γ if the chosen minimum degree is smaller than the true x_{\min} . $\hat{\gamma}$ is acceptable if the value of minimum degree is chosen being slightly higher than the true x_{\min} , because the decrease of sample number will slow down the deviation (Clauset *et al.*, 2009).

The estimation of x_{\min} is embedded in the Kolmogorov-Smirnov (KS) technique which is the common method that seeks to identify the maximum distance between the two sets of non-normally distributed data (Clauset *et al.*, 2009):
$$D = \max_{x \geq x_{\min}} |S(x) - P(x)| \quad (3.63),$$
 where $S(x)$ represents the complimentary cumulative degree distribution function (CCDF) of observed data satisfying $x \geq x_{\min}$, and $P(x)$ is the CCDF for the desired power law model at $x \geq x_{\min}$. The estimated x_{\min} is therefore corresponding to the value that minimizes D .

Goodness-of-fit test

Goodness-of-fit tests can be conducted to test the hypothesis that the observed data set follows the power law (Clauset *et al.*, 2009). These tests generate p values that indicate the plausibility of the power law fits. If p is close to 1, the data set is more likely to obey the power law distribution. In contrast, if $p < 0.1$, it is considered that the power law fit is moderate fit. If p closes to zero such as 0.004 (around 0.00), the hypothesis of power law distribution can be rejected.

According to Clauset *et al.* (2009), the procedures of calculating p start by measuring KS that is the distance between the empirical data and the hypothesized power law model. Thereafter, a large number of data sets are generated with the given scaling parameter and x_{\min} embedded in the original hypothesized

power law data set. KS^* of each generated data set is calculated relative to the best power law fit of that generated data. P is the fraction of KS^* s that is larger than KS in the set of KS^* .

Likelihood ratio test

Although results generated from the KS test indicate that the data distribution plausibly follows the power law, the data set could fit equally good or better in an exponential or a log-normal distribution (Clauset *et al.*, 2009). The likelihood ratio test is conducted to compare two candidate distributions. The likelihood ratio R is positive if the data fit is more likely to be the first distribution, and is negative if the data distribution is more likely to be close to the second distribution (Alstott *et al.*, 2014).

$R = \sum_{i=1}^n [\ln p_1(x_i) - \ln p_2(x_i)]$ (3.64), where $\ln p_j(x_i)$, $j=1,2$, represents i^{th} log likelihood value within distribution p_j . The significance of R value is p , which is calculated embedded in the normalized R :

$p = \left| \int_{R/\sqrt{2n}\sigma}^{\infty} \exp\left(\frac{-t^2}{2}\right) dt \right|$ (3.65), where σ is the root of expected variance of a single term, σ^2 , with

$\sigma^2 = \frac{1}{n} \sum_{i=1}^n \left[(\ell_i^{(1)} - \ell_i^{(2)}) - \left(\frac{1}{n} \sum_{i=1}^n \ell_i^{(1)} - \frac{1}{n} \sum_{i=1}^n \ell_i^{(2)} \right) \right]$ (3.66), $\ell_i^{(j)} = \ln p_j(x_i)$, and $\sqrt{n}R / \sigma$ is the normalized R .

3.6.3.2 Identifying critical mass members

A centrality measure seeks to identify members that play an important role in the global structure of networks (Shi, 2011). There are a set of centrality measurements but with different focus of recognizing the importance of members. For instance, eigenvector indices and PageRank consider the neighbours of neighbours who can connect with many others that are important. In this study, the degree centrality measure is adopted; because critical mass members are initial contributors (Oliver and Marwell, 1988) who may win more and more connections over time. In other words, the importance of one member is measured by the numbers of the nearest neighbours. Assuming a network with N nodes, the maximum

linkages toward a node is $N-1$. The degree centrality of node i with k degrees is calculated through (e.g.

Wang *et al.*, 2006): $DC_i = \frac{k_i}{N-1}$ (3.67).

3.6.3.3 The critical point

The dynamic network can be denoted as $G(t) = [N(t), L(t), f(t) : R]$, which is a time-varying 3-dimensional composition with $N(t)$ representing a set of nodes evolving over time, $L(t)$ indicating the developing of linkages/edges between nodes, and $f(t)$ demonstrating the function how nodes are associated, dropped or rewired with respect to a set of rules R that can be either/both internal or/and external influential factors on $f(t)$ (Lewis, 2011). A dynamic network may be either convergent- reaching its final state within a finite time, or divergent- changing forever with a manner of progressing infinitely. In the context that one can only predict the evolution of online forums in infinite time, this study follows the suggestions by Lewis (2011) to study the emergence of networks by observing the arising of a new phenomenon in a finite number of applications of micro rules. In this study, it is defined that t^* is the threshold point when the online network is developed from launch to a stable state, i.e. connected.

Far before the critical threshold point t^* , an online forum is disconnected with small clusters. Close to the critical threshold point t^* , the spinning cluster is emerging (Molloy and Reed, 1995; Newman *et al.*, 2001; Shi, 2011). The spinning cluster involves members who own majority connections within a network, and are therefore those who play an important role in information communication of that network (Newman *et al.*, 2001; Shi, 2011). This definition can be understood as the synonym of critical mass members who support the payoff of information contributions within online communities (Wasco *et al.*, 2009). Above t^* , a spinning cluster within which there is always a path that connects any of two members within the network is presented. Spinning cluster is associated with infinite system (Cohen *et al.*, 2003; Newman *et al.*, 2006); therefore, the presence of spinning cluster can predict a continuous growing of online forum. In fact, as online forum is scale-free, the network is continuously growing and developing into a stable state (spinning cluster always exists) (Albert and Barabasi, 1999). The critical

mass members' influence on the continuous evolution of a network (Olivier and Marwell, 1989) is therefore reflected.

For a random network, the condition for the emergence of giant component is measured by the ratio of the second movement against the first movement that should be at least 2 (Molloy and Reed, 1995; Cohen *et al.*, 2003). Denoting $\langle k^2 \rangle$ represents the second movement (second order of neighbours or the neighbours of neighbours), and $\langle k \rangle$ is the first movement (first order of neighbours or the nearest neighbours). If loop (self-connection) can be ignored, Cohen *et al.* (2003) argue that for a member who belongs to the spinning cluster, this member should have at least one outgoing edge that connects that member to others. Thus, the probability of member i in the spinning cluster is also connected with member j can be written

as (Cohen *et al.*, 2003): $\langle k_i | i \leftrightarrow j \rangle = \sum_{k_i} k_i P(k_i | i \leftrightarrow j) = 2$ or $\frac{\langle k^2 \rangle}{\langle k \rangle} = 2$ (3.68).

Phase transition is associated with percolation theory which calculates the critical point (t^*) to break down a network by randomly removing a fraction p of members (Cohen *et al.*, 2003; Newman, 2005). Incorporating (2.26), (2.27) and (2.28) (see Chapter 2), a fraction of $1-p$ of member remains but with a

new degree distribution (Cohen *et al.*, 2003):
$$\frac{\langle k_0^2 \rangle (1-p)^2 + \langle k_0 \rangle p(1-p)}{\langle k_0 \rangle (1-p)} = 2$$
 (3.69).

(3.69) can be reduced to $1-p = \frac{\langle k_0 \rangle}{\langle k_0^2 \rangle - \langle k_0 \rangle} = \frac{\langle k_0 \rangle}{\langle k_0(k_0-1) \rangle}$ (3.70). p is the critical point (t^*) after which spinning cluster occurs.

For the scale-free network, the degree distribution is: $P(k) \sim ck^{-\lambda}$, with $k_{\min} < k < k_{\max}$. The ration of the

second movement and the first movement can be (Cohen *et al.*, 2003): $\left| \frac{2-\gamma}{3-\gamma} \right| k_{\min}^{\gamma-2} k_{\max}^{3-\gamma}$, for $2 < \gamma < 3$.

As k_{\max} is the order of N , it diverges when $N \rightarrow \infty$. Thus, theoretically there is no critical threshold point for the scale-free network. However, for any finite network, phase transition can be observed although $(1-$

p) where p is closing to 1. In this study, results show that p is around 0.008, before which the online forum is disconnected and after which spinning cluster presents. Exact at the critical point (t^*), the size of critical mass members (largest cluster) scales as (Cohen *et al.*, 2003): $N^{\frac{\gamma-2}{\gamma-1}}$, in this study, it suggests a small group of 53 contributors within a network of 147190 members (around 0.04%).

3.6.4 Data visualisation software and analysis language

This study seeks to visualise the structure of online forum networks and describes the structural characteristics of non-linearity. After having compared different software to visualize the structural dynamics of online forums, the software Gephi is used. This is able to analyse exploratory data collected from the real world and to review dynamic structure evolving over time (<http://gephi.github.io>).

Also, the Python language is used for data analysis. The source codes for this computer language are openly shared from the Internet such as Stack Overflow and Google code (which was moved to Github recently) where programming hobbyists and enthusiasts frequently seek answers provided by others. In addition, the Python official website provides rich documentation with codes that could satisfy the different requirements by searching the key words.

Chapter 4 Results and discussions: study one – examinations of causal factors associated with online voluntary contribution of knowledge

4.1 Introduction

Study one seeks to understand how the key identified antecedents act together and influence knowledge contribution behaviours in the context of online forums. Data is collected through the online survey. Following the multistage data analysis discussed in chapter 3, the results of the pilot test are firstly analysed in order to examine the feasibility of measurement items adopted in this study. Modifications are made embedded in the results of the pilot test in terms of validity and reliability of constructs.

The main study follows similar procedures of the pilot test; in addition by the end the structural equation modelling is developed in order to test the hypothesised relationships within the conceptual model in chapter 2. The following sections will present results of the pilot test and the main study sequentially.

4.2 Pilot test

The pilot test is performed after the pre-test for evaluating the sense of the questionnaire. The pilot test of the online survey involves samples of members of online forums, and uses a paper-based survey form for collecting data over three weeks. It is designed in order to evaluate the questions and the reliability and validity of the multi- item scales adopted in this study. A descriptive data analysis is firstly performed before conducting the two-step approach CB-SEM.

Initially 80 cases are collected during the first day, which is used for the pilot test. As discussed in chapter 3, the response rates to this online survey are very high without missing cases.

Multi-scales are used in this study. Cronbach's alpha (α) is performed in order to examine the internal consistence that can reflect the reliability of the measurement scales (e.g. Hair *et al.*, 2010). Although Cronbach's alpha is not sufficient for conducting SEM, it is a commonly used reliability test on the measurement and scales that can give a first insight into the reliability of the questionnaires (e.g. Cortina, 1993). Table 13 summarizes the alpha test.

Table 13: Alpha test

	Cronbach Alpha	Numbers of items
Intention	0.735	4
Attitude	0.859	5
Perceived behavioural control	0.726	4
Subjective norms	0.764	4
Perceived critical mass	0.812	12
Trust in online communities	0.866	8
Trust in members	0.882	9

A value of Cronbach's alpha over 0.7 suggests a good internal consistency, and a value over 0.6 refers to an acceptable level (Cortina, 1993). However, a greater number of items will increase artificially the value of Cronbach's alpha and vice versa (Cortina, 1993). For this stage, the final questionnaire is confirmed.

4.3 The main study

The main study uses the whole sample size including both the earlier and later responses. The data collection lasts three weeks resulting in 910 responses in total. The 460 responders who answer the questionnaire during the first half period are the earlier responders. The 450 responders who answer the survey during the second half period of data collection are later responders. Preliminary data analysis is associated with the descriptive data analysis, and is firstly conducted. This refers to dealing with outlier, data distribution and sample size effect. Having cleaned the data, the next stage is to create the covariance-based structural models (CB-SEM) embedded in the two-stage processing (Wright *et al.*, 2012). That is, structural models that seek to test the conceptual model in chapter 2 are examined after the development of the measurement models. Detailed analysis processes are explained in the following sections.

4.3.1 Preliminary data analysis

4.3.1.1 Missing data and outliers

As mentioned before, responses from the online survey intermediaries are completed without blank cases; missing data analysis is not performed in this study. The univariate outlier detection generates the standardised z scores for each variable. Hair *et al.* (2010) argue that outliers are identified if their z scores are greater than 4. There are few cases identified as outliers. However, removing outliers, the entire of associated population will be excluded from analysis (e.g. Hair *et al.*, 2010). In this study, outliers are not removed from the data analysis as they represent different views of opinions. It is normal that respondents will have different views. For instance, respondents who rate a very low score on “intention to contribute” but a very high score on “perceived critical mass” and “trust in members” are more likely to be the free-riders who want to benefit knowledge supplied by others, and it is consistent with the issue of public goods, as discussed in chapter 2.

4.3.1.2 Data distribution

The normality test of data distribution should be considered seriously, because many parametric statistical tests such as t-tests, analysis of variance and regression are embedded in the assumption of normal distribution of data (Gauss, 1777–1855). In this study, data is considered normally distributed following the results generated through kurtosis (peaking) and skewness (position) tests, whose values suggest that neither kurtosis nor skewness exists (see table 15).

The central limit theorem suggests that z-scores can be one criterion of the normality test (e.g. Hair *et al.*, 2006). For a small sample size (little more than 40), the accepted range of z-scores for kurtosis and skewness is ± 1.96 , which suggests that data is normally distributed at the significant levels of 0.05. For a sample of several hundred (>200), the criteria is settled to ± 2.58 , indicating the normality of data distribution at significant levels of 0.01 (Hair *et al.*, 2011; Ghasemi and Zahediasl, 2012). As a result of which, the kurtosis and skewness scores should be near zero for the normally distributed data.

Although using this method can minimise the consequences of non-normal distribute dataset, in particular when the sample size is large (Tabachnick and Fidell, 2007), the maximum likelihood estimated techniques adopted in SEM can tolerate the violation of non-normal distribution of data (Hair *et al.*, 2010). The data used in this study is considered normally distributed with both skewness and kurtosis at the acceptable level (near zero) (Ghasemi and Zahediasl, 2012).

4.3.1.3 Descriptive data analysis

With the dataset screened and cleaned, the descriptive data analysis is performed in order to have a general view on the survey on both individual variables and constructs (Hair *et al.*, 2010). Table 14 shows the sample profile, and table 15 summarizes the descriptive characteristics of data related to the individual variables.

Table 14: Sample profile

Variables		Training (Frequency)	Testing(Frequency)	Total (%)
Gender	Male	222	156	41.5%
	Female	238	294	58.5%
Age	< 20	2	7	1%
	21-35	207	302	55.9%
	36-50	159	113	29.9%
	>50	92	28	13.2%
Education	Secondary School	18	18	4%
	College	122	152	30.1%
	B.A	280	240	57.1%
	Master	40	34	8.1%
	Phd	0	6	0.7%
	Total	460	450	100%

58.5% of respondents are female. The majority of participants (55.9%) in the questionnaire are between 21 – 35 years old cites using online forums. A clear majority (57.1%) of them have a bachelor diploma. The sample was randomly separated into two groups: a training dataset (460 cases) and a testing dataset (450 cases). This refers to the use of a cross-validation technique to gauge the reliability of the measurement model (e.g. Gerbing and Hamilton, 1996).

Table 15: Descriptive data of individual variable

	Mean	Standard errors of mean	Standard deviation	Minimum	Maximum	Skewness	Kurtosis
INT1	3.72	0.022	0.649	2	5	0.187	-0.5
INT2	3.72	0.022	0.666	2	5	-0.032	-0.224
INT3	3.56	0.024	0.736	1	5	-0.185	-0.155
INT4	3.09	0.029	0.874	1	5	-0.025	-0.43
ATT1	3.63	0.026	0.785	1	5	0.393	-0.519
ATT2	3.62	0.026	0.778	2	5	0.369	-0.674
ATT3	3.56	0.024	0.727	1	5	0.247	-0.277
ATT4	3.49	0.025	0.749	1	5	0.503	-0.126
ATT5	3.57	0.026	0.789	2	5	0.294	-0.546
PCMB1	3.68	0.025	0.740	2	5	-0.202	-0.195
PCMB2	3.64	0.025	0.741	1	5	-0.292	-0.049
PCMBC1	3.92	0.023	0.693	1	5	-0.349	0.545
PCMBC2	3.75	0.024	0.734	1	5	-0.293	0.12
PCMG1	3.44	0.028	0.851	1	5	-0.288	-0.66
PCMG2	3.45	0.027	0.800	1	5	-0.184	-0.256
PCMLINK1	3.62	0.022	0.663	1	5	-0.058	-0.05
PCMLINK2	3.75	0.024	0.733	2	5	-0.206	-0.176
PCMLINK3	3.70	0.023	0.696	1	5	-0.149	0.092
PCMD1	3.54	0.023	0.691	2	5	-0.03	-0.221
PCMD2	3.51	0.023	0.697	2	5	-0.006	-0.231
PCMD3	3.68	0.023	0.702	1	5	-0.018	-0.066
TRCA1	3.55	0.023	0.682	1	5	0.065	0.001
TRCA2	3.80	0.021	0.643	2	5	0.081	0.162
TRCA3	3.58	0.023	0.643	2	5	-0.082	-0.123
TRCB1	3.69	0.023	0.697	1	5	-0.253	-0.012
TRCB2	3.55	0.023	0.682	1	5	-0.142	0.008
TRCI1	3.67	0.023	0.679	2	5	-0.09	-0.197
TRCI2	3.64	0.021	0.683	1	5	-0.013	-0.127
TRCI3	3.62	0.023	0.640	1	5	-0.141	0.04
TRMA1	3.57	0.022	0.680	1	5	-0.005	0.014
TRMA2	3.52	0.024	0.666	1	5	0.003	0.047
TRMA3	3.42	0.023	0.721	1	5	-0.139	-0.055
TRMB1	3.62	0.023	0.698	1	5	-0.098	-0.061
TRMB2	3.62	0.024	0.692	1	5	-0.067	-0.088
TRMB3	3.75	0.022	0.718	1	5	-0.208	0.142
TRMI1	3.80	0.022	0.664	2	5	-0.194	0.008
TRMI2	3.46	0.025	0.676	2	5	-0.298	0.175
TRMI3	3.61	0.022	0.742	1	5	-0.038	-0.233

	Mean	Standard errors of mean	Standard deviation	Minimum	Maximum	Skewness	Kurtosis
SN1	3.61	0.022	0.668	1	5	-0.364	0.151
SN2	3.56	0.023	0.668	1	5	-0.364	0.151
SN3	3.59	0.022	0.691	2	5	-0.052	-0.214
SN4	3.65	0.021	0.669	2	5	-0.073	-0.185
PBC1	3.65	0.022	0.642	2	5	0.024	-0.263
PBC2	3.57	0.023	0.665	2	5	0.083	-0.315
PBC3	3.48	0.023	0.687	2	5	0.151	-0.29
PBC4	3.57	0.023	0.665	2	5	0.083	-0.315

(Note: INT~ Intention; ATT~ Attitude; PCM~ Perceived critical mass; TRC ~ Trust in online forums; TRM~ Trust in members; SN~ Subjective norms; PBC~ Perceived behavioural control; see appendix1 “questionnaire” for further details.)

The first construct “intention” involved with intention to contributing knowledge online is measured through four items with 5 point Likert-scale where 1 represents “strongly disagree” and 5 means “strongly agree”. Around 46% of responses are “agree” in term of intention to contribute and around 38% of responses voting for “neutral”. The mean score for the construct of intention is around 3.58 with item 4 having the lowest mean score at 3.09. If the item 4 is removed from analysis, the overall mean score will increase.

The mean score for the construct “attitude” is 3.56, with the lowest mean score at 3.49 measured by item 4. Around half of (49 %) of respondents mention that they are “neutral” in terms of attitude of contribution online, 47.71% of responders agree to contribute online. Only 3.48% of responders have a negative attitude of contribution to online discussions.

The mean score of the construct “perceived behavioural control” is 3.70, with the item “PBC3” having the lowest mean score at 3.48, and the highest “PBC1” at 3.65. 46.37% of responses chose the “agree”, and 42.20% for “neutral”. It is noted that 0% strongly disagree that both the conditions as well as self-efficiency favour contribution of knowledge online.

The overall mean scores of the construct “subjective norms” are 3.32. The majority of responders (49.96%) chose “agree” on items for measuring the perceived subjective norms. 0% of respondents rate each item of “subjective norms” at the “strongly disagree” level. 39.71% are at the level of “neuter”, and 6.74% at “strongly agree”.

Table 16 Descriptive data of constructs (in percentage %)

Construct	Strongly disagree	Disagree	Neuter	Agree	Strongly agree	Mean
IN	0.60	8.21	37.64	45.66	7.88	3.58
ATT	0.13	3.45	48.70	34.37	13.34	3.56
PBC	0.00	4.18	42.20	46.37	7.25	3.70
SN	0.00	3.59	39.71	49.96	6.74	3.32
TRC	0.11	3.09	37.36	51.22	8.22	3.32
TRM	0.18	4.88	38.36	49.02	7.56	3.83
PCM	0.14	5.93	34.66	50.63	8.64	3.65

(Note: INT~ Intention; ATT~ Attitude; PCM~ Perceived critical mass; TRC ~ Trust in online forums; TRM~ Trust in members; SN~ Subjective norms; PBC~ Perceived behavioural control; see appendix1 “questionnaire” for further details.)

The construct “perceived critical mass” has five dimensions measured through in total twelve items. Results suggest an average score for “perceived critical mass” of 3.65, with the item CMBC1 the highest at 3.92. Variances between the mean scores between items are not huge. When asking their opinions on the perceived critical mass within online forums, the majority of responders (50.63%) chose the options “agree”, and 34.66% for “neutral”. It is noted that 0.14% of responses strongly disagree, 8.64% strongly agree on items for perceived critical mass emerging online. This may indicate that the perceived critical mass members can be one antecedent of intention to discuss online.

The overall mean score for the construct “trust in online community” is 3.32, which is measured through three dimensions (ability, benevolence and integrity) representing eight items. Item “TRCA2” has the highest mean score at 3.8. In contrast, the lowest mean scores -3.55- are represented by the items

“TRCA1” and “TRCB2”. 59.44% of responses trust each item that is employed to measure the trust in online communities. 37.36% of responders are neutral with regard to trusting online forums.

The mean score of the construct of trust in members is around 3.83. The item “TRMI1” has the highest mean score at 3.8, and the item “TRMA3” has the lowest at 3.42. Similar to the descriptive data for measuring the constructs trust in online forums, most respondents rate their opinions on “trust in members” at the “agree” level (57.27%). 34.66% at the “neutral” level, and only 0.14% at the “strongly disagree” level.

In summary, the majority of respondent rate either “agree” or “neutral” for all items adopted in this study. The proportion of the extreme values represented by “strongly disagree” or “strongly agree” is less than 10%.

4.3.1.4 Non-responses analysis

The non-responses analysis is conducted by comparing the responses from the earlier responders and the later responders. The data collection lasted three weeks, responders who handed over their questionnaires during the first half of the three weeks are defined as the earlier responders, and the rest of the responders are named as the later responders.

The earlier respondents count comprised 50.5% of the total sample, of which 58.7% is male and 44.7% female. Most respondents are between 21-50 years old (85.8%), and a few respondents are under 20 years old (1%). It suggests that people who are between 20 to 50 years old are more likely to participate in activities within online forums. Given the fact that the population between 20 and 50 years old make up the majority of the whole population in the world, it therefore happens that individuals who are classified into this age category are more likely to be selected with the random sampling techniques. In addition, most participants have education at the college level (57.1%), with few (0.7%) at PhD level. This is again consistent with the general knowledge in terms of education level. As a result of discussions above, the sample frame is valid for further analysis.

Armstrong and Overton (1977) recommend the independent t-test as the technique to compare the mean differences between two groups. In this study, the independent t-test is performed to compare whether there is difference between the earlier responders and later responders groups with respective to all constructs (intention, subjective norms, perceived behaviour control, trust in online forum, trust in members and perceived critical mass). The following table illustrates results by performing independent t-test on all constructs.

The results produced by the Levene's test for equality of variances suggest that there are equal variances for the two groups. Although items such as “ATT1” have p-values which are less than 0.05, which indicates that differences exist between groups, the calculated effect sizes through the formula illustrated in chapter 3 (3.8) (effect size for t-value= $\frac{t^2}{t^2 + (N-1)}$) assume the small effect size (< 0.06) (Cohen, 1988). That is, at 95% significant level, the null hypothesis that assumes there is no differences between the two groups (earlier responders and later responders) in terms of the mean scores ($\mu_1 = \mu_2$) remains for all constructs. The non-response errors are therefore minimised (Armstrong and Overton, 1977).

Table 17: Independent t-test: early and late respondents

		Levene' test for equality of variances		t-test for equality of means							Effect size
		F	Sig.	T	DDL	Sig. (two-tailed)	Different means	Different standard error	Confidential interval 95 %		(t*t/(t*t+N-1))
									lower	upper	
INT1	H0	.018	.894	-2.210	908	.027	-.095	.043	-.179	-.011	
	H1			-2.208	900.905	.027	-.095	.043	-.179	-.011	
INT2	H0	2.620	.106	-1.254	908	.210	-.055	.044	-.142	.031	
	H1			-1.252	896.138	.211	-.055	.044	-.142	.031	
INT3	H0	.201	.654	.019	908	.985	.001	.049	-.095	.097	
	H1			.019	906.810	.985	.001	.049	-.095	.097	
INT4	H0	2.012	.156	-.603	908	.547	-.035	.058	-.149	.079	
	H1			-.603	905.957	.547	-.035	.058	-.149	.079	
ATT1	H0	8.651	.003	-.054	908	.957	-.003	.052	-.105	.099	small

		Levene' test for equality of variances		t-test for equality of means							Effect size
		F	Sig.	T	DDL	Sig. (two- tailed)	Different means	Different standard error	Confidential interval 95 %		(t*t/(t*N-1))
									lower	upper	
	H1			-.054	894.941	.957	-.003	.052	-.105	.100	
ATT2	H0	8.131	.004	-.690	908	.490	-.036	.052	-.137	.066	small
	H1			-.689	897.385	.491	-.036	.052	-.137	.066	
ATT3	H0	5.295	.022	-2.546	908	.011	-.122	.048	-.217	-.028	small
	H1			-2.543	894.671	.011	-.122	.048	-.217	-.028	
ATT4	H0	8.485	.004	-1.502	908	.133	-.074	.050	-.172	.023	small
	H1			-1.500	893.431	.134	-.074	.050	-.172	.023	
ATT5	H0	8.834	.003	-1.290	908	.197	-.067	.052	-.170	.035	Small
	H1			-1.289	894.292	.198	-.067	.052	-.170	.035	
PCMB1	H0	.031	.861	-.437	908	.662	-.021	.049	-.118	.075	
	H1			-.437	906.415	.662	-.021	.049	-.118	.075	
PCMB2	H0	.138	.710	-1.092	908	.275	-.054	.049	-.150	.043	
	H1			-1.092	905.682	.275	-.054	.049	-.150	.043	
PCMBC1	H0	.136	.712	-2.071	908	.039	-.095	.046	-.185	-.005	
	H1			-2.069	892.449	.039	-.095	.046	-.185	-.005	
PCMBC2	H0	1.077	.300	-.249	908	.803	-.012	.049	-.108	.083	
	H1			-.249	900.901	.804	-.012	.049	-.108	.084	
PCMG1	H0	4.195	.041	1.935	908	.053	.109	.056	-.002	.220	Small
	H1			1,934	900,681	.053	.109	.056	-.002	.220	
PCMG2	H0	5.502	.019	-.891	908	.373	-.047	.053	-.151	.057	Small
	H1			-.892	904.282	.373	-.047	.053	-.151	.057	
PCMLINK1	H0	.013	.908	.489	908	.625	.021	.044	-.065	.108	
	H1			.489	907.284	.625	.021	.044	-.065	.108	
PCMLINK2	H0	6.684	.010	.838	908	.402	.041	.049	-.055	.136	Small
	H1			.837	896.473	.403	.041	.049	-.055	.136	
PCMLINK3	H0	.668	.414	1.606	908	.109	.074	.046	-.016	.165	
	H1			1.606	907.995	.109	.074	.046	-.016	.165	
PCMD1	H0	.361	.548	-2.967	908	.003	-.136	.046	-.227	-.046	
	H1			-2.964	894.740	.003	-.136	.046	-.227	-.046	
PCMD2	H0	.384	.536	1.806	908	.071	.083	.046	-.007	.172	
	H1			1.805	906.921	.071	.083	.046	-.007	.172	
PCMD3	H0	.065	.798	-1,194	908	.233	-.055	.046	-.146	.035	
	H1			-1.194	907.856	.233	-.055	.046	-.146	.035	

		Levene' test for equality of variances		t-test for equality of means							Effect size
		F	Sig.	T	DDL	Sig. (two- tailed)	Different means	Different standard error	Confidential interval 95 %		(t*t/(t*N-1))
									lower	upper	
TRCA1	H0	.006	.941	-.841	908	.401	-.039	.047	-.130	.052	
	H1			-.840	906.815	.401	-.039	.047	-.131	.052	
TRCA2	H0	.000	.983	1.189	908	.235	.054	.045	-.035	.143	
	H1			1.190	908.000	.235	.054	.045	-.035	.142	
TRCA3	H0	.919	.338	-1.545	908	.123	-.066	.043	-.149	.018	
	H1			-1.544	895.134	.123	-.066	.043	-.150	.018	
TRCB1	H0	2.446	.118	1.294	908	.196	.060	.046	-.031	.150	
	H1			1.293	904.215	.196	.060	.046	-.031	.150	
TRCB2	H0	8.788	.003	.491	908	.623	.022	.045	-.067	.111	Small
	H1			.491	892.608	.624	.022	.045	-.067	.111	
TRCI1	H0	.853	.356	1.687	908	.092	.076	.045	-.012	.164	
	H1			1.687	905.841	.092	.076	.045	-.012	.164	
TRCI2	H0	2.753	.097	1.084	908	.279	.049	.045	-.040	.138	
	H1			1,083	905,705	.279	.049	.045	-,040	,138	
TRCI3	H0	.553	.457	-.124	908	.901	-.005	.042	-.089	.078	
	H1			-.124	904.451	.901	-.005	.042	-.089	.078	
TRMA1	H0	.413	.521	.624	908	.533	.028	.045	-.060	.117	
	H1			.624	904.915	.533	.028	.045	-.060	.117	
TRMA2	H0	.329	.566	.067	908	.947	.003	.044	-.084	.090	
	H1			.067	905.983	.947	.003	.044	-.084	.090	
TRMA3	H0	.141	.707	1.093	908	.275	.052	.048	-.042	.146	
	H1			1.093	907.939	.275	.052	.048	-.042	.146	
TRMB1	H0	6.183	.013	-.008	908	.993	.000	.046	-.090	.090	Small
	H1			-.008	891.135	.993	.000	.046	-.091	.090	
TRMB2	H0	.057	.811	-.515	908	.607	-.025	.048	-.118	.069	
	H1			-.515	907.916	.607	-.025	.048	-.118	.069	
TRMB3	H0	.059	.808	.772	908	.440	.034	.044	-.052	.120	
	H1			.772	907,708	.440	.034	.044	-.052	.120	
TRMI1	H0	.155	.694	.777	908	.437	.038	.049	-.058	.135	
	H1			.777	906.136	.438	.038	.049	-.058	.135	
TRMI2	H0	.119	.730	.886	908	.376	.039	.044	-.048	.126	
	H1			.886	907.547	.376	.039	.044	-.048	.126	
TRMI3	H0	.723	.395	2.011	908	.045	.101	.050	.002	.200	
	H1			2.010	903.375	.045	.101	.050	.002	.200	
SN1	H0	2.887	.090	1,749	908	.081	.080	.046	-.010	.170	
	H1			1.748	901.116	.081	.080	.046	-.010	.170	

		Levene' test for equality of variances		t-test for equality of means							Effect size	
		F	Sig.	T	DDL	Sig. (two- tailed)	Different means	Different standard error	Confidential interval 95 %	(t*t/(t*t+N-1))		
								lower	upper			
SN2	H0	1.504	.220	.997	908	.319	.044	.044	-.043	.131		
	H1			.997	903.534	.319	.044	.044	-.043	.131		
SN3	H0	4.137	.042	2.876	908	.004	.122	.042	.039	.205	Small	
	H1			2.876	907.252	.004	.122	.042	.039	.205		
SN4	H0	.195	.659	-.394	908	.693	-.018	.045	-.106	.070		
	H1			-.394	902.896	.694	-.018	.045	-.106	.070		
PBC1	H0	8.826	.003	.721	908	.471	.032	.044	-.055	.118	Small	
	H1			.721	895.609	.471	.032	.044	-.055	.119		
PBC2	H0	3.867	.050	-.322	908	.747	-.015	.046	-.104	.075		
	H1			-.322	898.239	.748	-.015	.046	-.104	.075		
PBC3	H0	.323	.570	-.510	908	.610	-.024	.047	-.115	.068		
	H1			-.510	907.993	.610	-.024	.047	-.115	.068		
PBC4	H0	.723	.395	2.011	908	.045	.101	.050	.002	.200		
	H1			2.010	903.375	.045	.101	.050	.002	.200		

INT~ Intention; ATT~ Attitude; PCM~ Perceived critical mass; TRC ~ Trust in online forums; TRM~ Trust in members; SN~ Subjective norms; PBC~ Perceived behavioural control; see appendix1 “questionnaire” for further details.

4.3.1.5 Comparison of groups of respondents

Before hypothesis testing, the dataset is evaluated through the independent sample t-tests in terms of the dichotomy categorical variable (sex) and analysis variance test (ANOVA) for control variables (age and education). This analysis seeks to understand whether the hypothesis results could be generalised by the evaluation of no significant differences occurring in responses to all variables within the online survey (Hair *et al.*, 2010).

Results of one-way ANOVA show there is no significant difference between group means (see tables 18-20). Although respondents with different backgrounds with respect to gender, age and education have demonstrated few differences in rating on items for several constructs in this survey, the variances of differences are less than 2%. For instance, female responders may be more likely to trust others and

online forums than male responders do, but that differences (<1%) can be neglected in analysis. Responders over 50 years old may be more likely to demonstrate their trust in others (Mean=3.72), but with less than 2% of change that can be counted by the treatment. Again, there is less than 2% of change in group means that responders with a PhD diploma may be more likely to have a positive attitude and higher self-efficiency toward online discussions. In other words, there is the possibility of using one structural CB-SEM model in order to generalize findings.

Table 18: Independent t-test: comparing means between genders

		Mean	F(p)	T	Effect size
Intention	M	3.55	H0: 1.648(0.200)	1.46	
	F	3.50		1,45	
Attitude	M	3.56	H0: 3.089(0.079)	-0.52	
	F	3.58		-0.52	
Perceived behavioural control	M	3.57	H0: 0.527(0.468)	0.00	
	F	3.57		0.00	
Subjective norms	M	3.59	H0: 2.981(0.085)	-0.48	
	F	3.61		-0.48	
Trust in members	M	3.55	H0:5.514(0.019)	-1.91	0.004(small)
	F	3.61		-1.89	
Trust in online forums	M	3.61	H0: 4.043(0.045)	-1.81	0.001(small)
	F	3.67		-1.78	
Perceived critical mass	M	3.62	H0 :13.711(0.000)	-0.84	
	F	3.65		-0.82	

Table 19: One-way ANOVA: comparing means between ages

	<20 (mean)	21-35 (mean)	36-50 (mean)	>50 (mean)	F-value (p-value)	Eta ²
Intention	3.58	3.54	3.48	3.54	0.782(0.504)	
Attitude	3.56	3.59	3.53	3.63	0.947(0.417)	
Perceived behavioural control	3.70	3.54	3.59	3.62	1.079(0.357)	
Subjective norms	3.32	3.57	3.58	3.72	2.025(0.109)	
Trust in members	3.32	3.57	3.58	3.72	4.580(0.003)	0.015(small)
Trust in online forums	3.83	3.63	3.65	3.68	0.796(0.498)	
Perceived critical mass	3.65	3.62	3.64	3.69	1.003(0.391)	

significant difference at $p < 0.05$; the sum squares for the construct “trust in members” are:3.047(within group) and 200.914(between group)

Table 20: One-way ANOVA: comparing means between education levels

	High school (mean)	Some College (mean)	College (mean)	Master (mean)	PhD (mean)	F-value (p-value)	<i>Eta</i> ²
Intention	3.50	3.52	3.50	3.68	3.50	1.854(0.117)	
Attitude	3.55	3.48	3.62	3.54	3.63	2.615(0.034)	0.011(small)
Perceived behavioural control	3.65	3.53	3.56	3.73	3.78	2.410(0.048)	0.011(small)
Subjective norms	3.53	3.60	3.60	3.61	3.83	0.426(0.791)	
Trust in members	3.53	3.60	3.58	3.64	3.56	0.375(0.827)	
Trust in online forums	3.74	3.61	3.64	3.73	3.81	1.359(0.246)	
Perceived critical mass	3.59	3.65	3.62	3.66	3.77	0.441(0.779)	

significant difference at $p < 0.05$; the sum squares for the construct “attitude” are : 3.90(within group) and 337.337 (between groups); the sum squares for the construct “perceived behavioural control” are: 2.903(within group) and 272.908(between groups)

4.3.2 Measurement validity

The preliminary data analysis seeks to clean and screen the dataset and validate the sample. The descriptive data analysis illustrates a summarization of responses. Cross-validation technique is applied to gauge the reliability of the measurement model (Gerbing and Hamilton 1996). The sample is randomly separated into two groups: a training dataset (460 cases) and a testing dataset (450 cases) (note: the training and testing datasets are not necessarily the earlier and later responses that are divided chronologically).

Principal component analysis (PCA) is used to assess face validity (Hair *et al.* 2010) and is conducted on the training dataset. Results are cross-validated with the testing dataset where the confirmatory factor analysis (CFA) is performed. The full sample is used for examining the final measurement model and for establishing the structural model. CFA is undertaken following the process of covariance-based structural equation modelling (CB-SEM) (Wright *et al.*, 2012). CB-SEM is preferred when the model involves multi-dimensional constructs on a large sample size (Wright *et al.*, 2012). Online trust and perceived critical mass are assumed as multi-dimensional constructs, and the sample size is sufficient using the

ad hoc rule of thumb of a 10:1 ratio of sample size to the number of free parameters (Kahai and Cooper, 2003), which would require around 600 responses.

4.3.2.1 Exploratory factor analysis

The principle component analysis is the technique to reduce a number of variables into a smaller number of artificial components, which takes into account the variances in the raw data (e.g. Hair *et al.*, 2010). In this study, the principle component analysis is the extraction method in order to examine the eigenvalues of items adopted from previous research.

Principle component analysis extraction with an oblique rotation method is performed in this study. Different from varimax rotation method, the oblique rotation is often undertaken when the correlated factors are a plausible representation of reality (Browne, 2001). Embedded in the cut-off point of 0.5 (Hair *et al.*, 2010), variables that do not reach this cut-off in terms of communalities and factor loadings are removed.

The sample adequacy is examined through KMO and Bartlett's test. The Kaiser-Meyer-Olkin measure indicates that the sample size is suitable for PCA, with a score of 0.957 closing to 1 and exceeding 0.5. The Bartlett's test of sphericity suggests a significant value of 0.000, indicating that statistically significant relationships between variables (Hair *et al.*, 2010).

Table 21: indices KMO and Bartlett test

Indices of Kaiser-Meyer-Olkin for the quality of sample size		0.957
Bartlett's test of sphericity	Khi-deux approx.	11099.102
	DDL	990
	Signification	0.000

Only the observations for the earlier responders are used for analysis

Seven components are extracted from the 46 observed variables embedded in Kaiser's criteria, and explain 57.83% of total variance (>50%). Each component (factor) has an eigenvalue greater than 1.

Variables with communalities and factor loadings less than 0.5 are removed from analysis. The step is repeated to ensure all communalities and factor loadings are satisfactory. The total variance extracted is improved to 67.27% after removing unsatisfactory items.

It is noted that there remains 5 items for the construct “trust in online community”, “attitude” and “perceived critical mass”. However, the limitation of PCA is that it is difficult to identify the latent variables (Hair *et al.*, 2010), which can be compromised with CFA explained in the following section. The following tables illustrate 17 variables excluded after PCA and 29 variables remained for further analysis sequentially. Finally the face validity of the seven factors (constructs) is established.

Table 22: Results of PCA

Constructs	Items remained	Items removed	Cronbach alpha
Intention	IN1		0.780
	IN2		
	IN3		
		IN4	
Attitude	ATT1		0.809
	ATT2		
	ATT3		
	ATT4		
	ATT5		
Perceived behavioral control	PBC1		0.852
	PBC2		
	PBC3		
		PBC4	
Subjective norms	SN1		0.838
	SN2		
	SN3		
Trust in online forums	TRCA1		0.835
	TRCA2		
	TRCB2		
	TRCI1		
	TRCI3		
		TRCA3	
		TRCI2	
		TRCB1	
Trust in members	TRMA2		0.751
	TRMA3		
	TRMB1		
	TRMI2		
		TRMA1	
		TRMB2	
		TRMB3	
		TRMI1	
		TRMI3	

Constructs	Items remained	Items removed	Cronbach alpha
Perceived critical mass			0.717
	PCMLINK1		
	PCMLINK2		
	PCMD1		
	PCMD2		
	PCMD3		
		PCMB1	
		PCMB2	
		PCMG1	
		PCMG2	
		PCMBC1	
		PCMBC2	
		PCMLINK3	

INT~ Intention; ATT~ Attitude; PCM~ Perceived critical mass; TRC ~ Trust in online forums; TRM~ Trust in members; SN~ Subjective norms; PBC~ Perceived behavioural control; see the appendix 1 “questionnaire” for further details.

4.3.2.2 Testing reliability

CFA is the main method for evaluating the factors’ validity and reliability as explained in chapter 3. However, Cronbach’s alpha (α) is initially tested for minimizing the potential and uncontrolled errors in terms of the variables’ reliability.

Results of alpha (α) tests indicate the measurement scales on the seven constructs are reliable, with all scores over 0.7. The following section will discuss the data analysis by performing a CFA method to test the validity and reliability of factors and the observed variables loading to factors.

4.3.2.3 Confirmatory factor analysis

Without issues such as missing and non-normal distribution to address, results generated from EFA are cross validated through CFA. CFA is performed on the proposed measurement model. In SEM, factors are understood as the latent constructs which are measured through the observed variables. The maximum likelihood estimation technique is employed in CFA. The following are the explanation of results of CFA representing the same sequential explanations in chapter 3.

Model one hypothesizes that all observed indicators load on a first order factor. The bad model fit indices of model one suggest that the observed variables are multidimensional and further inform that the second order factors can be formed. Model two assigns the predictors' loadings to the respective first order factors. If the assumption of multidimensionality for the first order model is supported, model two produces better fit indices than model one. Convergent validity is assessed by the standardized factor loadings that are higher than 0.5 at 0.001 significant level (two tailed) (Tabachnick and Fidell 2007; Hair *et al.* 2010). Model three tests the discriminant validity between paired factors and is supported if there are significant changes of chi-squared values between the models representing with and without correlated factors (Zait and Berteau, 2011). The second order factor is introduced in models four–six. The parallel model treats each dimension equally by restricting both the factor loadings and residual variances. The tau model releases the residual variances but restricts the equal loadings, thereby allowing different means for each item. The congeneric model frees all constraints in the parallel model, thereby assuming different means and variances to each item (Graham, 2006). Comparing the model fit indices generated from the parallel, tau and congeneric models, the best fit model is considered as the reliability estimations (Wright *et al.*, 2012).

Model one: first-order factor model

Within the conceptual model, there are three multidimensional constructs, “trust in online forums”, “trust in members” and “perceived critical mass”. The first model hypothesizes that the first-order factor accounts for the variances of all 15 items remaining in the three multidimensional constructs. The results prove a poor model fit with chi-square (421.090), d.f (90), CFI (0.879), and RMSEA (0.091). It suggests that indicators do not load on to one factor.

Model two: convergent validity

In this model, the 15 indicators specify 3 freely correlated first order latent variables. Results show an improved model fit compared with model one: chi-square (689.298), d.f (187), CFI (0.926), and RMSEA

(0.072), which supports that the multidimensional model, rather than a single factor model, is a plausible representation of reality. Standardised factor loadings to their specified factors are all over 0.5 but with a majority less than 0.7; they all are significant at 0.001 levels ($p < 0.001$). The convergent validity is therefore supported (Wright *et al.*, 2010).

It is noted that factors are highly correlated (0.9 between trust in members and trust in online forums; 0.89 between trust in online forums and perceived critical mass), which may suggest that some predictors cross loading toward different factors. The values of squared multiple correlations (squared R) for each observed variable are therefore examined. The squared R explains the variances of the exogenous

variables counted by the endogenous variables (e.g. Bentler and Raykov, 1998) :

$$R^2 = 1 - \frac{\hat{\sigma}_{residual}^2}{\hat{\sigma}_{endogenous}^2} \quad (4.1),$$

where $\hat{\sigma}_{residual}^2$ is the estimated variance of residuals, and $\hat{\sigma}_{endogenous}^2$ is the implied (estimated) variance of endogenous variables. The value of the squared R can therefore be inferred as the predictability of exogenous variables within a latent variable structural equation model (Bentler and Raykov, 1998), or the reliability of predictors. Predictors with squared R values less than 0.30 are removed, which suggests that the predictors can explain less than 30% of variance counted by the latent variable.

The scores of standardised factor loadings improved with a majority over 0.7 and all significant at 0.001 levels. Although the overall model fit indices suggest a better model fit, with chi-square = 327.277(80), CFI=0.955, and RMSEA =0.058(CI 90%: 0.052, 0.065), the correlation between trust in online forums and perceived critical mass (0.85) is still high.

It is not uncommon in marketing research using CB-SEM when struggling to fit adjustments to eliminate meaningful items (Hair *et al.*, 2010; Hair *et al.*, 2014). Hair *et al.* (2014) argue that the construct context should be weighted over the model fit adjustment. For this stage, results show that the convergent validity of observed variables loading toward their specified latent variables is meaningful, and the study moves on. The following highlights the results with observed variables remaining for the specified latent variables.

Model three: discriminant validity

Discriminant validity is evaluated by comparing chi-square differences of pair of factors, where different methods are available. One most rigorous method is to set unconstrained (uncorrelated) and constrained models (correlated) (Bagozzi *et al.*, 1991). If the chi-squared of the constrained model is significantly bigger than that of the unconstrained model, the hypothesis of discriminant validity between paired constructs remains (Wright *et al.*, 2012). In AMOS, the unconstrained model is created by setting “1” to the variances of latent variables, and releasing the “1” constraint from the observed to latent variable. The constrained model is to set “1” to the correlation between paired constructs. The discriminant validity test can be performed comparing the hierarchic chi-squared value or model fit indices between the constrained and unconstrained model. Results show the evidence of discriminant validity for all paired factors.

Table 23: discriminant validity

<i>Factor</i>	<i>Unconstrained model χ^2 (df)</i>	<i>Constrained model χ^2 (df)</i>	<i>P</i>
Trust in members vers:			
Attitude	62.034(17)	562.438(18)	0.000
Trust in online community	460.033(14)	994.830(15)	0.000
Perceived critical mass	51.176(8)	530.622(9)	0.000
Perceived behavioral control	51.176(8)	530.662(9)	0.000
Subjective norms	51.406(8)	500.060(9)	0.000
Intention	40.289(8)	605.401(9)	0.000
Trust in online community vers:			
Attitude	185.755(34)	313.897(35)	0.000
Perceived critical mass	39.076(8)	81.616(9)	0.000
Perceived behavioural control	113.236(19)	113.822(20)	0.000
Subjective norms	62.440(19)	346.513(20)	0.000
Intention	68.619(19)	208.669(20)	0.000
Perceived critical mass vers :			
Attitude	154.263(19)	260.407(20)	0.000
Perceived behavioural control	62.784(19)	166.247(20)	0.000
Subjective norms	101.065(19)	296.884(20)	0.000
Intention	63.149(19)	241.766(20)	0.000
Attitude vers :			

<i>Factor</i>	<i>Unconstrained model χ^2(df)</i>	<i>Constrained model χ^2(df)</i>	<i>P</i>
Perceived behavioural control	78.015(19)	211.357(20)	0.000
Subjective norms	101.065(19)	296.884(20)	0.000
Intention	103.209(19)	332.249(20)	0.000
Perceived behavioural control vers :			
Subjective norms	22.536(8)	74.238(9)	0.004
Intention	61.306(8)	149.410(9)	0.000
Subjective norms vers :			
Intention	20.051(8)	233.359(9)	0.000

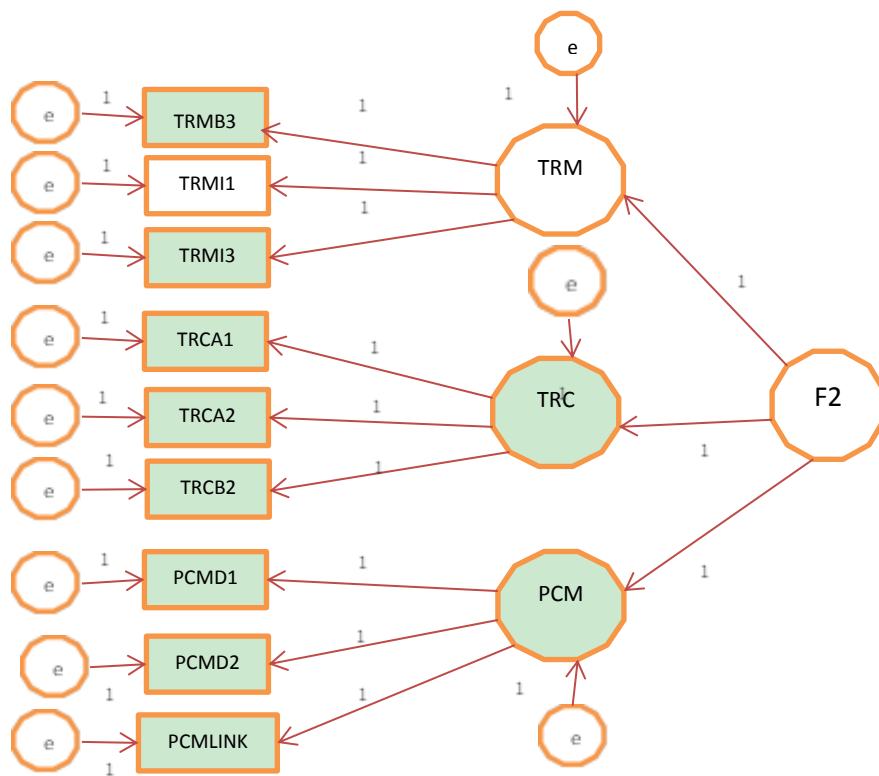
Model four: parallel model

As a reminder, the covariance-based model involves the second order factor that is the cause of relationships between the first order factors. The parallel model also named path model restricts factor loadings and residual variances to be equal in order to treat each dimension as being equal (Wright *et al.*, 2010).

Results indicate a poor model fit which suggests that the first order factors are not equally representing the second order factor (Wright *et al.*, 2012). However, it may happen because the consequences of the first order factors are represented in different degrees within the conceptual model. For example, “trust in members” positively influence on “subjective norms”, and “attitude”, but “perceived critical mass” only positively effects on “subjective norms”. As a consequence of this, the first order factor may not equally reliably represent the second order factors since the vector of error terms associated with “trust in members” may vary according to the factors to which it is leading.

To understand the accuracy and equality of the first order factors, the model fit indices of the parallel model should be compared with those of the tau equivalent model and congeneric model (Wright *et al.*, 2012). Table 24 summarizes results generated from the parallel model, the tau equivalent model and the congeneric model.

Figure 12: the initial parallel model



INT~ Intention; ATT~ Attitude; PCM~ Perceived critical mass; TRC ~ Trust in online forums; TRM~ Trust in members; SN~ Subjective norms; PBC~ Perceived behavioural control; see the appendix 1 “questionnaire” for further details.

Model Five: Tau equivalent

The tau equivalent model allows the residual components to differ. It still assumes that the true scores of the latent variables are the same. Results suggest that the first order factors are well modelled.

Model six: congeneric model

With the congeneric model, constraints of equal factor loading and residual variance settled in the parallel model are released. Results show that the assumption of equal relationships between indicators is not realistic for the data collected for the online survey, and that indicators are not parallel. The chi-square differences between the parallel model and the tau equivalent model suggest that the first order factors vary in representing the second order factor; the chi-square differences between the tau equivalent model

and the congeneric model show that the first order factors are influenced by the second order factors in different ways (Wright *et al.*, 2012). Thus, the congeneric model is selected as the more reliable model.

Table 24: model fit indices

	χ^2	df	CFI	RMSEA	RMSEA 90% C.I
Parallel model	3103.455	42	0.000	0.327	(0.322, 0.332)
Tau equivalent	96.604	30	0.954	0.069	(0.054, 0.085)
Congeneric	64.554	24	0.971	0.061	(0.044, 0.076)

df = degrees of freedom; GFI = Goodness-of-Fit Index; RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index; PCFI = Parsimony Goodness of Fit Index; C.T. = Confidential Intervals

Testing the measurement model with full samples

It is recommended to test the final measurement model using the full sample (n= 910) in order to generalize results from the measurement model (Gerbing and Hamilton, 1996). The process is repeated in terms of construct reliability and validity. The overall model fit indices of the congeneric model suggest the best model fit, with results illustrated in table 25. The final measurement model is summarized in table 26.

Table 25: overall model fit indices

	χ^2	df	CFI	RMSEA	RMSEA 90% C.I
Measurement model	132.350	24	0.959	0.070	(0.059, 0.079)

df = degrees of freedom; GFI = Goodness-of-Fit Index; RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index; PCFI = Parsimony Goodness of Fit Index; C.T. = Confidential Intervals

Table 26 : final measurement model

Constructs	Items	Standardized factor Loadings	p-values	a
Intention (Erden <i>et al.</i> , 2012)	I try to share knowledge with online forums members. (INT1)	0.83	***	0.780
	I plan to share knowledge with online forums members. (INT2)	0.85	***	

Constructs	Items	Standardized factor Loadings	p-values	a
	I openly share information that I gained from news, magazines and journals with other online forums members. (INT3)	0.64	***	
Attitude (Ajzen, 1991)	For me, sharing my knowledge with other members is pleasant. (ATT1)	0.75	***	0.809
	For me, sharing my knowledge with other members is enjoyable. (ATT2)	0.76	***	
	For me, sharing my knowledge with other members is beneficial. (ATT3)	0.69	***	
	For me, sharing my knowledge with other members is good. (ATT4)	0.72	***	
	For me, sharing my knowledge with other members is valuable. (ATT)	0.76	***	
Perceived behavioural control (Reagans and Mcevily, 2003)	If I want, I always could share knowledge with online forums members. (PBC1)	0.74	***	0.852
	It is always possible for me to share my knowledge with network members. (PBC2)	0.74	***	
	I enjoy giving my true opinion, which is not risky. (PBC3)	0.59	***	
Subjective norms (Yen <i>et al.</i> , 2011)	There is a high level of cooperation (e.g. replying to other members' questions and comments) among members of the online forum. (SN1)	0.71	***	0.838
	Members are willing to sacrifice time and effort for the benefit of this online forum. (SN2)	0.78	***	
	There is a high level of sharing among members of this online forum. (SN3)	0.68	***	
Trust in online forums (Ridings <i>et al.</i> , 2002)	My forum is very competent. (TRCA1)	0.76	***	0.835
	My forum is able to satisfy its members. (TRCA2)	0.72	***	
	If a member required help, my forum's members would do their best to help. (TRCB2)	0.61	***	
Trust in members (Mcknight and Chervany, 2002)	Members throw their hearts into the communities' affairs. (TRMI1)	0.72	***	0.751
	Member's suggestions are the best they can offer. (TRMI3)	0.61	***	
	Members will help each other solve problems. (TRMB3)	0.63	***	

Constructs	Items	Standardized factor Loadings	p-values	a
Perceived critical mass (Sledgianowski and Kulviwat, 2009)	Many people participate in the discussions. (PCMD1)	0.71	***	0.717
	Many of my friends participate in the discussions. (PCMD2)	0.74	***	
	I know the member(s) who give valuable suggestions, and they become my online friend(s). (PCMLINK2)	0.62	***	

a = Cronbach's alpha; *** = significant at 0.001

4.3.3 Developing the structural model

The measurement model is examined in terms of convergent and discriminant validity (e.g. Wright *et al.*, 2012). The structural model examines the significance, strength, and the correlations between constructs within the conceptual model and nomological valid (e.g. Anderson and Gerbing, 1988). The examinations of the measurement and structural models represent the two-step approach of SEM (e.g. Anderson and Gerbing, 1988). Table 27 summarizes the hypotheses developed within the conceptual model. The original structural model is recalled in figure 20:

Figure 13: The original structural model

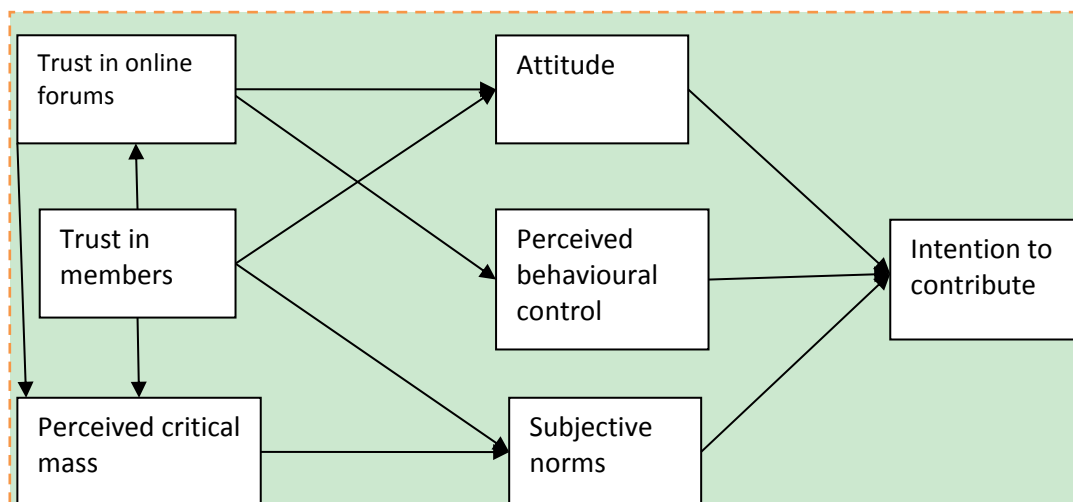


Table 27: hypotheses developed in the conceptual model

H1	Attitudinal beliefs have a positive influence on intention to continuous sharing knowledge online.
H2	PBC beliefs has a positive impact on members' continuous intention to share knowledge
H3	SN is positively associated with intention to share knowledge online.
H4a	Trust beliefs in members have a positive impact on subjective norms.
H4b	Trust beliefs in members have a positive impact on attitude.
H4c	Trust beliefs in members can lead to trust beliefs in online forums.
H5a	Trust beliefs in online forums have a positive impact on attitude.
H5b	Trust beliefs in online forums have a positive impact on perceived behavioural control.
H6a	Perceived critical mass beliefs have positive influence on SN.
H6b	Perceived critical mass beliefs have positive influence on trust in members.
H6c	Perceived critical mass beliefs have positive influence on trust in online forums.

4.3.3.1 Model fit indices of structural model

The overall model identification and model fit are assessed before rejecting hypothesis (ese) or accepting new path(s). The structural model exhibits a good model fit as illustrated in the following table. With the sample size (910), the evaluation of χ^2 or adjusted χ^2 may not be adequate (Schumacker and Lomax, 2004), and other model fit indices are felt more appreciated. The model fit indices discussed in the chapter 3 show the requirements for how a good model fit is satisfied. The squared multiple correlation which equals to 0.571 for the factor “intention” to contribute online suggests that the structural model can predict around 57% (>50%) of variances (Hair *et al.*, 2010). It therefore suggests that the factors “trust in members” , “trust in online forums” and “perceived critical mass” are the key antecedents of the factors “attitude”, “perceived behavioural control” and “subjective norms” are good predictors to “intention” to contribute online.

Table 28: Model fit indices of the hypothesized model

	Hypothesized model	Cut off of good fit
Absolut fit indices		
χ^2	766.586	P<0.001
Df	216	Positive
χ^2/df	3.595	<5
GFI	0.929	>0.9
RMSEA	0.053 (C.I.90%:0.049,0.058)	<0.07 acceptable <0.05 very good
Incremental fit indices		
CFI	0.942	>0.9
Parsimonious fit indices		
PCFI	0.804	>0.5

df = degrees of freedom; GFI = Goodness-of-Fit Index; RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index; PCFI = Parsimony Goodness of Fit Index; C.T. = Confidential Intervals

The overall model fit and the variances extracted by the structural model are satisfied, the hypothesised relationships are assessed which is discussed in the following section.

4.3.3.2 Parameter estimates for the structural paths

The path estimations show the significant inter-construct relationships with the null hypothesis of a correlation between two variables being zero. Thus the alternative hypothesis assumes that their correlation is different from zero (either positive or negative). A path estimate which is significant at the 0.001 /0.05/0.01 levels (two-tailed) can suggest the direct effects of variable on subsequent variable. As a result of this, a significant of p-value in the hypothesis test demonstrates the relationships between two variables. Testing significance of individual path generates results that 9 of 11 paths are statistically significant at 0.001 (two tailed) and left 2 paths which are significant at 0.01 levels (two tailed).

Significance at 0.001 levels (two-tailed) suggests that 1 case out of 1000 may be different from the conclusion. Significance at 0.01 levels (two-tailed) provides a wider confidence interval, e.g. 90%, were true values on 90% of occasions. Results support DTPB and can be applied to understand online

contribution behaviours. However, it is noted that the subjective norms have less power in predicting intention to contribute online.

Table 29 Regression weights of the hypothesized model

Path	Estimate	S.E.	C.R.	P	Decision
Trust in online forums <--- Trust in members	.872	.057	15.171	***	Support
Perceived critical mass <--- Trust in members	-.079	.069	-1.151	.250	Not support
Perceived critical mass <--- Trust in online forums	.859	.073	11.739	***	Support
Subjective norms <--- Perceived critical mass	.411	.063	6.503	***	Support
Attitude <--- Trust in online forums	.866	.085	10.165	***	Support
Perceived behavioural control <--- Trust in online forums	.855	.043	19.732	***	Support
Subjective norms <--- Trust in members	.594	.068	8.679	***	Support
Attitude <--- Trust in members	-.065	.090	-.725	.469	Not support
Intention <--- Attitude	.370	.042	8.833	***	Support
Intention <--- Perceived behavioural control	.297	.067	4.423	***	Support
Intention <--- Subjective norms	.198	.058	3.447	***	Support

*p<0.05(two-tailed); **p<0.01(two-tailed);***p<0.001(two-tailed); S.E.: Standard Errors; C.R.: Critical ratio (>1.96 for the factor covariance being significant)

The majority of hypotheses developed in the conceptual model can be accepted, with statistical positive influences from the antecedents to consequences (p<0.001 two-tailed). However, it is suggested that trust in members has direct negative influence on attitudinal beliefs in online contribution (β = -0.065, p = 0.469, two-tailed), which is the opposite of the original hypothesis that trust in members can positively influence attitude to contribute online. This hypothesis is therefore rejected.

Darley and Latané (1968) are the first to discover the bystander effect in which individuals would not be likely to offer their help to victim(s) when others are present. The antecedents of the bystander effect are complex (not the purpose of this thesis), within which the diffusion of responsibility plays an important role in it (Darley and Latané, 1968). That is, individuals are likely to presume others will take actions so that they are not responsible for an event, and this is more likely to happen in groups of certain critical size and when the responsibility is not clearly resigned (Darley and Latané, 1968; Leary and Forsyth, 1987). With regard to the context of online forums, it is argued that one's trust in the ability and

benevolence of fellow members may decrease one's perception of urgency / valuable / beneficence in helping others. For instance, the more likely one trusts that others will post quality answers to questions; the less likely that one is urged to help in answering questions. These sorts of beliefs that maintain that answering questions is not urgent / valuable may have a negative influence on the attitude to contribute online.

4.3.3.3 Examining plausibility of models

The original hypothesized model has achieved a satisfactory model fit level as well as explained variances over 50%. In addition, it supports significant relationships between latent constructs. However, the original hypothesized structural model is over identified with $\chi^2=895.568(219)$ and $p<0.001$. It is possible that an alternative model can represent the dataset more efficiently. Following the suggestions from Weston and Gore (2006), the alternative model is firstly developed by deleting non-significant paths, i.e. the arrow from trust in members to attitude. The alternative model also evaluates modifications in explained variances. The second alternative model evaluates the variances explained through the model modification indices.

Removing un-supported path

With the original model, the hypothesis representing the positive influence of trust in members on attitude is not supported. By delating this path, results show a slightly better good model fit. χ^2 increases by 0.237 to 908.760 with a higher degree of freedom equals to 199 (>198). The χ^2/df ratio decrease by 0.022 to 4.567 and the parsimonious fit indices PCFI is increased of 0.004 to 0.794. Regarding to GF1, CFI and RSEMA indices, there is no change.

The squared R value for "intention" remains the same as 0.572. That is, the alternative model extracts the same variances as the original hypothesized model can do. In addition, estimates of paths from antecedents to their subsequent variables remain the same. Testing the path significances and directions show that all hypotheses can be accepted, with all being positively different from zero at 0.001 levels. In

summary, it is concluded that the alternative model one is better than the original model, because there is no issue concerning the path significance to address, and the comparative model fit indices (also parsimonious) PCFI indicate is slightly improved.

Table 30 Regression weights of alternative model one

Path		Estimate	S.E.	C.R.	P
Trust in online forums	<--- Trust in members	.861	.057	15.130	***
Perceived critical mass	<--- Trust in online forums	.796	.047	16.897	***
Subjective norms	<--- Perceived critical mass	.412	.066	6.277	***
Attitude	<--- Trust in online forums	.748	.047	15.827	***
Perceived behavioural control	<--- Trust in online forums	.855	.043	19.778	***
Subjective norms	<--- Trust in members	.588	.071	8.318	***
Intention	<--- Attitude	.423	.051	8.224	***
Intention	<--- Perceived behavioural control	.275	.067	4.121	***
Intention	<--- Subjective norms	.205	.057	3.625	***

*p<0.05(two-tailed); **p<0.01(two-tailed);***p<0.001(two-tailed) S.E.: Standard errors; C.R.: Critical ratio

Testing variance extracted

The modification indices proposed to convey the observed variables “Attitude3” and “Attitude4” (the highest M.I.:45.472). The overall model fit indices suggest a better model fit with χ^2 /df (4.106), GFI (0.918), CFI (0.930), RMSEA (0.058; C.I.:0.054, 0.062), and PCFI (0.081). The squared R for intention to contribute knowledge online is 0.571, with the same variances explained. However, the path significance examinations suggest that the hypotheses of trust in members to attitude and to perceived critical mass are not supported. The alternative model one is therefore chosen as the structural model that best explains the data set.

Table 31 Regression weights for alternative model two

Path			Estimate	P	Decision
PCM	<---	TRM	-0.081	0.238	Not supported
TRC	<---	TRM	0,325	***	Supported
PCM	<---	TRM	0 .668	***	Supported

Path			Estimate	P	Decision
TRC	<---	PCM	0,679	***	Supported
SN	<---	PCM	0,383	***	Supported
ATT	<---	TRC	0,763	***	Supported
PBC	<---	TRC	0,902	***	Supported
SN	<---	TRM	0,562	***	Supported
ATT	<---	TRM	-0,078	0.349	Not supported
IN	<---	ATT	0,406	***	Supported
IN	<---	PBC	0,274	***	Supported
IN	<---	SN	0,182	***	Supported

*p<0.05(two-tailed); **p<0.01(two-tailed) ;***p<0.001(two-tailed); PCM: perceived critical mass; TRM: trust in members;

TRC: trust in online forums; SN: subjective norms; ATT: attitude; IN: intention; PBC: perceived behavioural control

Table 32 Comparing model fit indices

Indices	Hypothesized model	Model 2	Model 1(chosen)
Absolut fit indices:			
χ^2	908.533	895.041	776.932
Df	198	218	219
χ^2/df	4.589	4.106	4.089
GFI	0.912	0.918	0.929
RMSEA	0.063 (C.I.90%:0.059,0.067)	0.058 (C.I.90%:0.054,0.062)	0.053 (C.I.90%:0.049,0.057)
Incremental fit indices:			
CFI	0.921	0.930	0.942
Parsimonious fit indices:			
PCFI	0.790	0.801	0.808

df = degrees of freedom; GFI = Goodness-of-Fit Index; RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index; PCFI = Parsimony Goodness of Fit Index; C.T. = Confidential Intervals

4.3.4 Testing mediated effects and moderated effects

A mediator could also be a moderator (Baron and Kenny, 1986). There are two main streams of previous studies on the combination of mediated and moderated effects, where the mediated and moderated effects are presented separately (Donaldson, 2001) or evaluated simultaneously (Edwards and Lambert, 2007;

Preacher *et al.*, 2007). The later often requires an experimental design and variables are measured in interval/ordinal levels, thus most of these design requirements are not satisfied by this study. For example, the mediated moderation model involves two predictors to the moderator which are embedded in the theoretical background. In this study, the moderated mediation effects can be tested, as it can be performed on continuous variables (Preacher *et al.*, 2007). The objective of this study is to create an integrated model that explains the important antecedents of intention to contribute online with respect to social and structural influences, without an experimental design.

A bootstrap method which re-samples 2000 times is conducted, with confidence interval set as 0.95. The analysis follows the logic proposed by Baron and Kenny (1986), and additionally uses structural equation modelling with latent variables (Muller *et al.*, 2005; Hopwood, 2007). An advantage of incorporating latent variables in contrast to observed variables is that it can ameliorate the reliability and method effects on the mediation and moderation models (Hopwood, 2007). The measurement errors associated with one particular observed variable that is specified loading toward the latent variable is unlikely to be shared with other observed variables that are equally loading toward the same latent variable, since latent variables seek to measure the overall desired effects (Hopwood, 2007).

The hypothesized model suggests that perceived critical mass has mediated/moderated effects on trust in members to subjective norms. The above examination of path significances using the CB-structural model also shows that the paths from trust in members to attitude and to perceived critical mass are not significant. However, the hypothesized model assumes that trust in online forums has mediated/moderated effects on (i) trust in members to perceived critical mass, and (ii) trust in members to attitude. The subsequent analyses seek to examine the causal relationships between the factors mentioned above, and extend results generated from the structural equation model using nested model where benefit factors are chosen, explained as follows.

Trust in members – perceived critical mass – subjective norms:

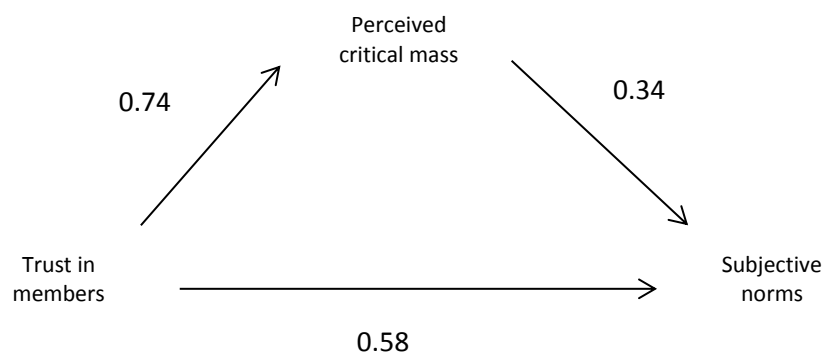
The mediated effects of perceived critical mass on subjective norms:

Step 1: subjective norms = $a_0 + C * (\text{Trust in members}) + \text{error}$;

Step 2 : subjective norms = $b_0 + C' * (\text{Trust in members}) + b_1 * (a_0 + a_1 * \text{Trust in members} + \text{error}) + \text{error}$;

Step 1 describes Y (subjective norms) which is regressed by X (trust in members) and the regression coefficient is C. The overall model fit indices are satisfied with $\chi^2(8) = 18.823$ ($p = 0.016 < 0.05$), CFI = 0.994 and RMSEA (0.039; C.I.90%:0.016, 0.062). The bootstrap results show that the direct effects of perceived critical mass on subjective norms are 0.661 (SE=0.047) and significantly different from zero at 0.001 levels (two-tailed; lower: 0.573; upper: 0.755). The first criterion of Baron and Kenny (1986) is met, suggesting that perceived critical mass is positively influence on subjective norms at 0.05 levels (one-tailed).

Figure 14: Mediation model one: the mediation effects of perceived critical mass on subjective norms



$$M = a_0 + a_1 X + \varepsilon ; \quad Y = b_0 + C' X + b_1 M + \varepsilon$$

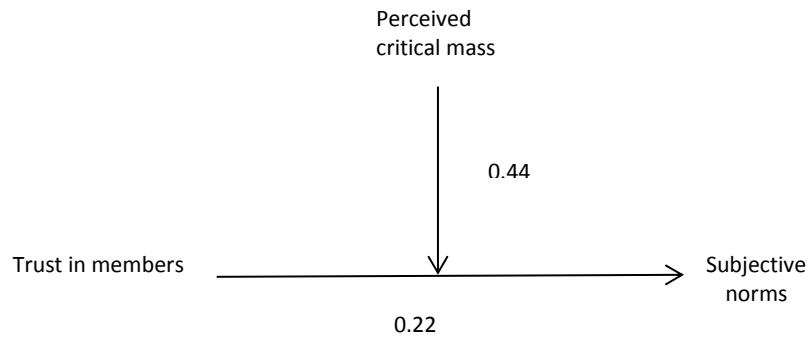
With the step 2, the mediator perceived critical mass is introduced, the overall model fit indices are equally satisfied, with $\chi^2(24) = 95.814$, CFI = 0.973 and RMSEA = 0.057 (C.I. 90%: 0.046, 0.070). The path from X (trust in members) to Me (perceived critical mass) is called path a. The path from Me to Y (subjective norms) is named path b. Both path a (0.74) and path b (0.34) are strong and significant at 0.001 levels (two-tailed). It is inferred therefore that the criteria 2 and 3 recommended by Baron and Kenny (1986) are satisfied.

The path coefficient from X to Y in the mediation model is named path c', which is different from zero at 0.001 significant levels. The indirect effects of trust in members on subjective norms through perceived critical mass are 31.5% ($=100\% * 1 - 100\% * c'/c$), and significantly different from zero at 0.001 levels. That is, the bootstrapping results suggests that the null hypotheses of the product of path a and path b equals to zero could be rejected with 95% of confidence. Because $\beta_{c'}$ is smaller than β_c , and is significant at 0.001 levels, it is inferred partial mediated effects of perceived critical mass on subjective norms (Baron and Kenny, 1986).

The overall model fit statistics for the moderator model two suggest a poor model fit, with $\chi^2(19) = 298.603$, CFI = 0.869 and RMSEA = 0.127 (C.I.90%: 0.115, 0.140). The effects of interaction of trust in members and perceived critical mass (TRM X PCM) on subjective norms are not statistically significant different from zero at 0.05 levels (C.R. = 1.176 < 1.96, $p = 0.086$), suggesting that there is no moderation effects of interaction of trust in members and perceived critical mass on subjective norms.

However, results are different from the structural model where the path from trust in members to perceived critical mass is not significant at 0.05 levels (path a in mediation model one). One possible explanation is that the mediation effects of perceived critical mass on trust in members are reduced with the introduction of other variables such as trust in online forums. In order to understand whether trust in online forums affects path a or path b or both (Fairchild and MacKinnon, 2009), three moderated mediation models are created following the recommendations of Preacher *et al.* (2007).

Figure 15: Moderation model two: the moderation effects of perceived critical mass on subjective norms



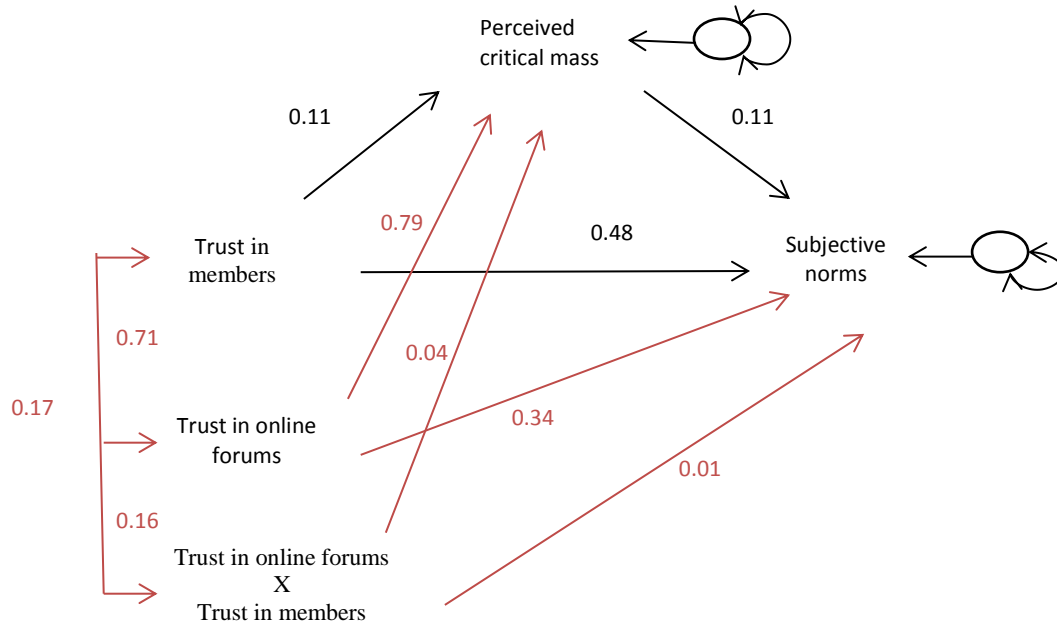
$$Y = a_0 + a_1X + a_2M_o + a_3XM + \varepsilon$$

Trust in online forums effects path a

The conditional indirect effects of X (trust in members) on Y (subjective norms) are expressed (Preacher *et al.*, 2007): $f(\hat{\theta} | M_o) = \hat{b}_1(\hat{a}_1 + \hat{a}_3M_o) = 0.11(0.11 + 0.040M_o)$, where M_o represents the moderator.

$\hat{\beta}_{a_3}$ (=0.040) is not significant at 0.05 levels (SE= 0.028, CR=1.362, p=0.173), suggesting that trust in online forums does not moderate path a. Results generated from bootstrapping show that the null hypotheses of the indirect effects (ab) are equal to zero cannot be rejected. There are no indirect effects of independent variables (trust in members) on dependant variable (subjective norms) through the mediator (perceived critical mass) with the introduction of trust in online forums as moderator on path a.

Figure 16: Moderated mediation model three: trust in online forums effects path a



$$Y = b_0 + C_1 'X + C_2 'M_o + C_3 'XM_o + b_1 M + \varepsilon$$

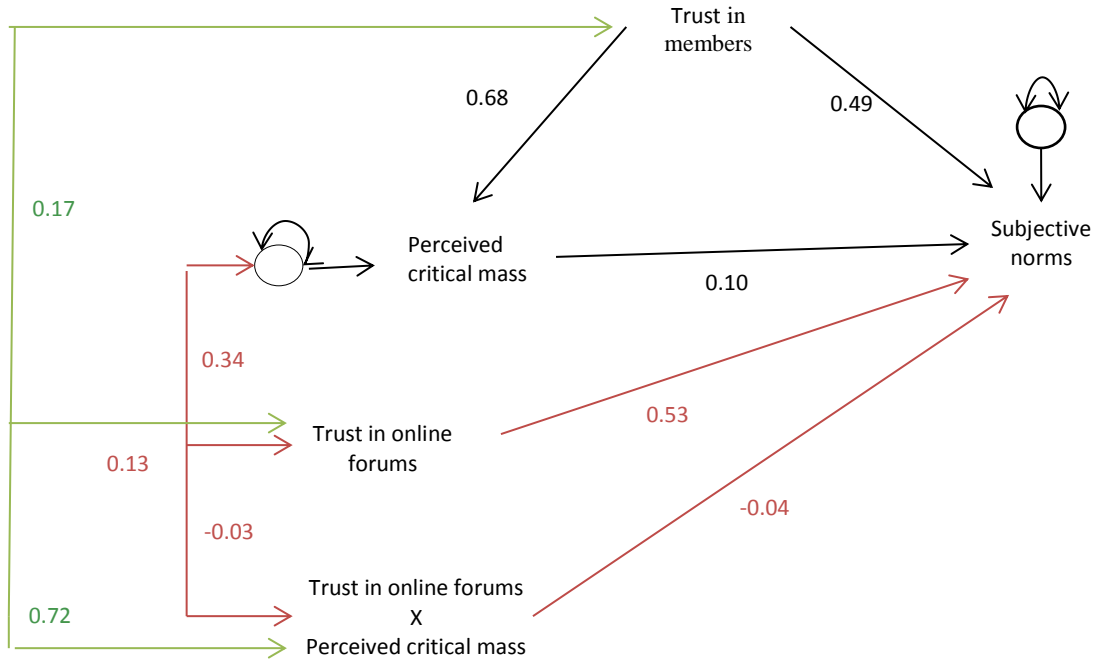
$$M = a_0 + a_1 X + a_2 M_o + a_3 XM_o + \varepsilon$$

$$M_o$$

Trust in online forums effects path b

With the moderated mediation model four, it is assumed that trust in online forums moderates the path from mediator (perceived critical mass) to outcomes (subjective norms). The conditional indirect effects of trust in members on subjective norms are (Preacher *et al.*, 2007): $f(\hat{\theta} | M_o) = \hat{a}_1(\hat{b}_1 + \hat{b}_3 M_o) = 0.68(0.10 - 0.04 M_o)$. $\hat{\beta}_{b_3} (= -0.042)$ is not significant at 0.05 levels (SE = 0.033, CR = -1.474, p=0.141), suggesting the b path is not moderated by the interaction of mediator and moderator. Bootstrapping tests show that perceived critical mass no longer mediates the relationship from trust in members to subjective norms, assuming trust in online forums moderates the path b.

Figure 17: Moderated mediation model four: trust in online forums effects path b



$$Y = b_0 + C_1 X + b_1 M + b_2 M_o + b_3 MM_o + \varepsilon$$

$$M = a_0 + a_1 X + \text{cov}(M, M_o) + \text{cov}(M, MM_o) + \varepsilon$$

Trust in online forums effects path a and path b

The moderated mediation model five assumes that trust in online forums moderates the mediation effects of perceived critical mass on subjective norms regressed on trust in members. The conditional indirect

effects within this model are expressed through (Preacher *et al.*, 2007): $f(\hat{\theta} | M_o) = (\hat{a}_1 + \hat{a}_3 M_o)(\hat{b}_1 + \hat{b}_2 M_o)$

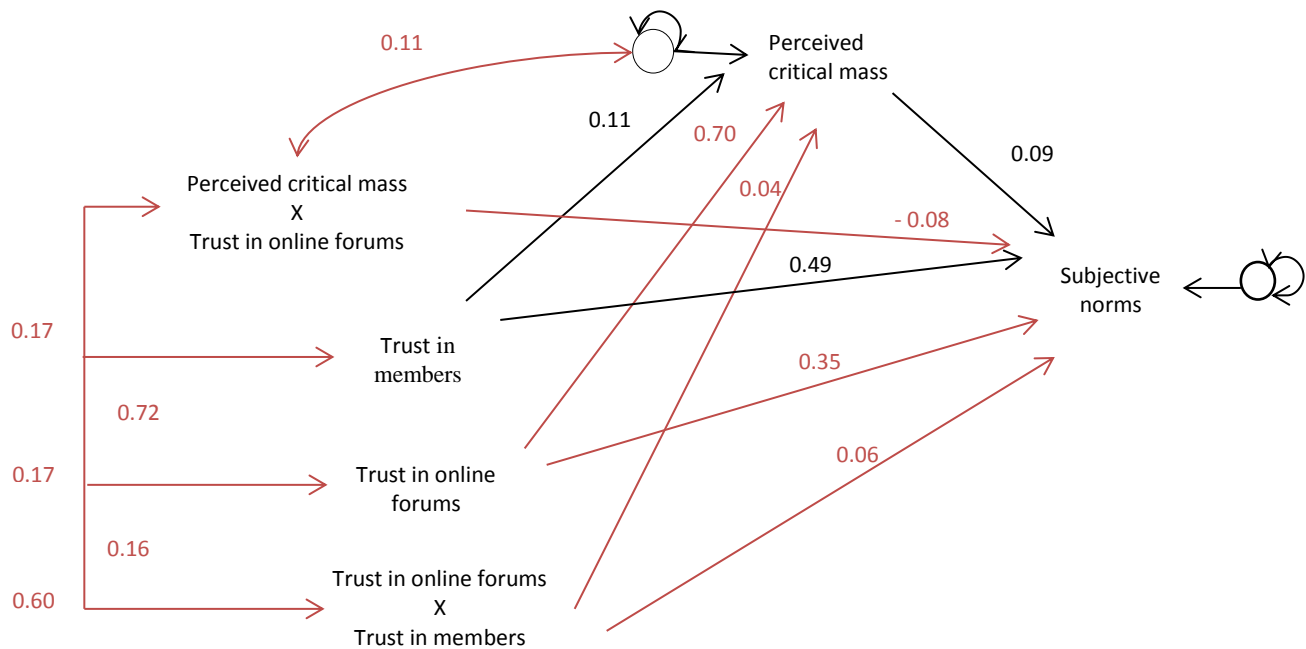
$= (0.11 + 0.040 M_o)(0.09 - 0.08 M_o)$. $\hat{\beta}_{c_3}$ ($= 0.06$) is not significant at 0.05 levels (SE=0.019, CR=1.616,

$p=0.106$), suggesting no significant moderation effects on the mediation assumptions. $\hat{\beta}_{b_2}$ ($= -0.08$) is

significant at 0.05 levels (SE=0.019, CR=-2.126, $p=0.033$), indicating a significant negative moderation

effects of trust in online forums on path b. $\hat{\beta}_{a_3}$ (=0.040) is not significant at 0.05 levels (SE=0.013, CR=1.351, p=0.177), suggesting there is no significant moderation effect on path a. A bootstrapping test suggests that there is no mediated effect of perceived critical mass on trust in members to subjective norms in the condition of trust in online forums moderating both path a and path b.

Figure 18: Moderated mediation model five: trust in online forums effects on path a and path b



$$Y = b_0 + b_1M + b_2MM_o + C'_1X + C'_2M_o + C'_3XM_o + \varepsilon$$

$$M = a_0 + a_1X + a_2M_o + a_3XM_o + \text{cov}(M, MM_o) + \varepsilon$$

Table 33: Moderation/Mediation effects of perceived critical mass

	Estimate	S.E.	C.R.	P	LOWER	UPPER
TRM→SN(c)	0.841	0.064	14.269	***	0.711	0.976
TRM→PCM(a)	0.744	0.059	13.131	***	0.626	0.866
PCM→SN(b)	0.335	0.066	6.073	***	0.192	0.458

	Estimate	S.E.	C.R.	P	LOWER	UPPER
TRM→SN(c')	0.576	0.086	8.547	***	0.415	0.754
PCMXTRM→SN	0.024	0.014	1.176	0.086	-0.022	0.051
TRCXTRM→PCM	0.04	0.028	1.362	0.173	-0.004	0.34
PCMXTRC→SN	-0.042	0.033	-1.474	0.141	-0.005	0.032
TRCXTRM→SN	0.06	0.019	1.616	0.106	-0.039	0.33

***significant at 0.001; *significant at 0.05; SN←PCMXTRM in moderation model two; TRCXTRM→PCM in moderated mediation model three; PCMXTRC→SN in moderated mediation model four; TRMXTRC→PCM in moderated mediation model five.

Table 34: Model fit indices: model N° one~five

	Chi-squares(df)	CFI	RMSEA	C.I 90%
Mediation model one	95.814(24)	0.973	0.057	(0.046,0.070)
Moderation model two	684.784(57)	0.845	0.078	(0.073,0.083)
Moderated mediation model three	226.997(56)	0.958	0.058	(0.050,0.066)
Moderated mediation model four	226.997(56)	0.958	0.058	(0.050,0.066)
Moderated mediation model five	241.583(64)	0.960	0.055	(0.048,0.063)

The above analyses suggest that perceived critical mass has partial mediated effects on trust in members (social influences) to subjective norms. However, such mediated effects of perceived critical mass disappear within the proposed integrative model. It is noted that trust in online forums does not moderate the mediated effects of perceived critical mass on the path from trust in members to subjective norms. However, it cannot exclude the possible influence of trust in online forums on the moderation effects of perceived critical mass, because the moderation effects of trust in online forums are not tested due to the lack of experimental design. It is also possible that there are (is) other(s) variable(s) which are not considered in the hypothesized model that can impact on the mediation effects of perceived critical mass on trust in members regress to subjective norms. However, it can be concluded that perceived critical mass has partial mediated effects on social factors.

Trust in members – perceived critical mass – trust in online forums

The mediated model six, which assumes perceived critical mass mediates the effects of trust in members on trust in online forums, benefits from good overall model fit indices. Trust in online forums completely mediates the effects of trust in members on perceived critical mass. The indirect effects, $\hat{\beta}_{ab}$ (0.561), are significantly different from zero, suggesting the mediated effects of trust in online forums on perceived critical mass. $\hat{\beta}_c$ (=0.670) is significant at 0.05 levels (SE=0.040, CR=11.234, $p < 0.05$), while $\hat{\beta}_{c'}$ (=0.115) is not significant at 0.05 levels (SE=0.072, CR=1.817, $p = 0.069$).

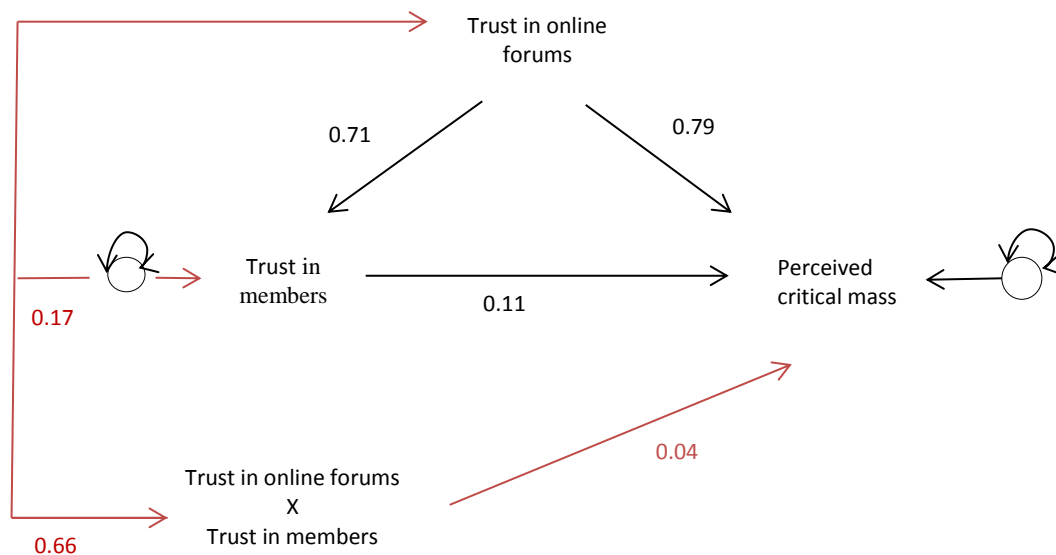
To understand whether the mediator (trust in online forums) is also a moderator, the moderation model seven is created to investigate the interaction of X (trust in members) and M_o (trust in online forums). The moderation model seven, which assumes trust in online forums moderates the effects of trust in members on perceived critical mass, suffers a poor model fit indices due to the collinearity between predictors, i.e. trust in members and trust in online forums. The high VIF (variance inflation factor) of the predictor will not have influence on the estimates but tends to increase the standard errors and p-values. However, collinearity can be tolerated with a large sample size as it is the case in this study. The VIF of suspicious constructs and their tolerances are computed. Results show that VIF for both trust in members and trust in online forums are less than 2.5, which is the recommended boundary for a high VIF by Allison (2014). As a result of this, the tolerances are close to 1 and the collinearity between trust in members and trust in online forums can be ignored at this stage.

Results generated from the moderation model seven suggest that trust in online forums has moderation effects on perceived critical mass, because the interaction of trust in members and trust in online forums regressed to perceived critical mass is statistically significant at the 0.05 level ($\hat{\beta}_{a_3} = 0.075$, SE= 0.030, C.R = 2.447, $p = 0.014$).

Table 35: *Collinearity test*

Model		Standardized	t	Sig.	Collinearity statistics	
		Bêta			Tolerance	VIF
1	(Constant)		.000	1.000		
	Trust in members	.186	6.445	.000	.726	1.378
	Trust in online forums	.557	19.344	.000	.726	1.378
2	(Constant)		-.900	.368		
	Trust in members	.180	6.217	.000	.719	1.391
	Trust in online forums	.553	19.196	.000	.722	1.384
	Trust in online forums X Trust in members	.054	2.164	.031	.972	1.029

Figure 19: Moderated mediation model eight: Trust in members effects on path b



$$Y = b_0 + b_1M + b_2XM + C'X + \varepsilon$$

$$M = a_0 + a_1X + cov(M, XM) + \varepsilon$$

Trust in online forums completely mediates the effects of trust in members on perceived critical mass, it is also the moderator of Y (perceived critical mass) regressed on X (trust in members). The magnitude levels of indirect effects could occur in different ways (Preacher *et al.*, 2007). In this case, the independent variable trust in members may act as a moderator and impact on the path from trust in online forums to perceived critical mass. The moderated mediation model eight is used to examine this idea. Results show that trust in members has no moderation effect on perceived critical mass regressed on trust

in online forums. This is demonstrated through the conditional indirect effects:

$f(\hat{\theta} | X) = \hat{a}_1(\hat{b}_1 + \hat{b}_2 X) = 0.71(0.79 + 0.040X)$. $\hat{\beta}_{b_2}$ ($=0.040$) is not significant because zero is setting within the 95% confidential interval ($p=0.178$). The null hypothesis of no conditional indirect effects cannot be rejected.

Table 36: Path estimations model N° six ~ eight

Relation	Estimates	S.E.	C.R.	P	LOWER	UPPER
PCM \leftarrow TRM(c)	0.670	0.040	11.234	***	0.587	0.747
PCM \leftarrow TRM (c')	0.115	0.072	1.817	0.069	-0.037	2.65
PCM \leftarrow TRC \leftarrow TRM	0.561	0.066	- -	- -	0.451	0.714
PCM \leftarrow TRMXTRC	0.075	0.030	2.447	**0.014	0.019	0.138
PCM \leftarrow TRCXTRM	0.040	0.028	1.346	0.178	-0.005	0.041

***significant at 0.001 **significant at 0.01 *significant at 0.05; PCM \leftarrow TRC \leftarrow TRM: the indirect effects in mediation model six; PCM \leftarrow TRMXTRC: in moderation model seven; PCM \leftarrow TRCXTRM: in moderated mediation model eight.

Table 37: Model fit indices: model N° six ~ eight

	Chi-squares(df)	CFI	RMSEA	C.I 90%
Mediation model six	212.543(24)	0.963	0.067	(0.055,0.079)
Moderation model seven	347.351(19)	0.837	0.138	(0.125,0.151)
Moderated mediation model eight	142.397(30)	0.958	0.064	(0.054,0.075)

As a summary of the above findings, trust in online forums completely mediates the effects of trust in members regressing to perceived critical mass. Moreover, it is plausible that as the level of trust in online forums increases, the perceived size of contributors within online forums increase. The moderation effects of trust in members and the mediation effects of trust in online forums on perceived critical mass are not found to occur simultaneously

Trust in members-trust in online forums-attitude

Trust in online forums completely mediates the effects of trust in members on attitude. Path c is significant at 0.001 levels ($\beta = 0.61$, C.R. = 10.678, $p < 0.001$), but path c' is no longer statistically significant with the mediator trust in online forums introduced in mediation model nine ($\beta = 0.051$, C.R. = 0.754, $p = 0.451$). The criteria proposed by Baron and Kenny (1986) are satisfied. The rational thinking is that there should not be an arrow from trust in members to attitude.

The moderation model ten disagrees that trust in online forums functions as moderator between trust in members and attitude. The interaction between the moderator and the independent variable regresses to the dependent variable is not significant ($\beta = 0.061$, C.R. = 1.933, $p = 0.053$). The overall model fit indices are not satisfied with the moderation model ten due to the collinearity between trust in members and trust in online forums but which can be tolerated as discussed above.

The moderated mediation model eleven, which seeks to understand the effects of X (trust in members) on path b, suggests that trust in members does not impact on the path b as the null hypothesis of $\hat{\beta}_{b_2} = 0$ (C.R. = 1.933, $p = 0.053$) cannot be rejected (not significant at 0.05 levels). This result is consistent with that generated through bootstrapping, because zero is between the lower (2.25 percentile) and upper boundary (97.25 percentile). The overall model fit indices of moderated mediation model eleven show a good model fit, suggesting little important values are lost in moving from completely freely estimated variance-covariance to the specified and restricted model (Hopwood, 2007).

Table 38: Mediation/moderation tests on the relationship between TRM-TRC-ATT

Relation	Estimate	S.E.	C.R.	P	LOWER	UPPER
ATT←TRM (path c)	0.61	0.057	10.678	***	0.507	0.717
ATT←TRM (path c')	0.051	0.085	0.754	0.451	-0.096	0.178
ATT←TRCXTRM	0.089	0.037	2.847	**0.004	0.007	0.153
ATT←TRMXTRC	0.061	0.038	1.933	0.053	-0.008	0.118

***significant at 0.001 **significant at 0.01*significant at 0.05 ATT←TRMXTRC: in moderation model ten;

ATT←TRCXTRM: in moderated mediation model eleven.

Table 39: Model fit indices: model N° nine ~ eleven

	Chi-squares(df)	CFI	RMSEA	C.I 90%
Mediation model nine	161.238(39)	0.969	0.059	(0.049,0.068)
Moderation model ten	471.916(32)	0.870	0.123	(0.113,0.133)
Moderated mediation model eleven	183.248(48)	0.966	0.056	(0.047,0.064)

Examination of the moderation and mediation effects can contribute to knowledge by clarifying the complex relationships between trust in members, trust in online forums and perceived critical mass. Previous studies have mainly tested the predictive regression relationships of individual variables, leaving little understanding of the causal relationships among them.

In summary, the examinations of the causal relationships between online trust and perceived critical mass are in agreement with the results by examining the structural model. In particular, the elimination of the path representing the effect of trust in members on attitude is embedded in the result that trust in online forums completely mediates and moderates the effects of trust in members on attitude, and the elimination of the path from trust in members to perceived critical mass is due to the influence of trust in online forums.

4.3.5 Summary of results

This study follows the two-step SEM process recommended by Anderson and Gerbing (1988). EFA is firstly conducted in order to reduce factor dimensions. Seven latent constructs are formed after EFA analysis. Following this, a covariance based measurement model (Wright *et al.*, 2012) successfully assessed construct validity and reliability. With the measurement model in place, the structural model is developed to test hypotheses proposed by the conceptual model.

Although the original hypothesized structural model demonstrates a good model fit, hypotheses about the positive influence of trust in members on attitude and trust in members on perceived critical mass are not

supported. Recommendations from Weston and Gore (2006) are adopted to evaluate other plausible structural models. In total two alternative models are developed with the alternative model one representing the elimination of unsupported paths, and model two evaluating variances extracted following the suggestions of model modification indices. After assessing the three structural models, it is concluded that the alternative model one with deleted unsupported paths is preferred, inferring all paths being significantly different from zero at the 0.001 level.

In order to understand why the path from trust in members to attitude is removed from the conceptual model, subsequent analyses are conducted to examine the causal relationships among proposed antecedents to intention to contribute online. This is undertaken using mediation, moderation and moderated mediation models. A bias-corrected percentile bootstrapping method is used to obtain the statistical power of these examinations. Latent variables are incorporated with regression analysis. The advantage of using latent variables can help to reduce measurement error (Muller *et al.*, 2005; Hopwood, 2007).

4.3.5.1 Discussion

This study has obtained results that confirm most hypotheses developed in the framework. Results agree with previous studies that interpersonal trust (Wasco and Faraj, 2005), institutional trust (Chen 2007; Erden *et al.*, 2012) and perceived critical mass (Shen *et al.*, 2013) are the key dynamic predictors of online knowledge contribution consistency. However, the identified predictors and their different magnitude influential levels have not been examined simultaneously with existed studies. With study one, trust in online forums is found having more weighted power in in the prediction of the determinants for online intentional contribution behaviours.

When taken individually, findings showed that trust in members has a positive impact on subjective norms (H7). This is in agreement with the arguments by Jeffries and Becker (2008) that high levels of trust in others may lead to social influences on intended behaviours. Trust in online forums acts as a contextual factor affecting the attitudinal and behavioural control beliefs (H4b and H6). These findings

may be related to results in previous studies. For instance, trust in online communities can lead to the positive attitude to purchase online (Zimmer *et al.* 2010); trust in the competences of online communities is positively associated with perceived behavioural control (Erden *et al.*, 2012). Given the fact that trust in online forums positively affects the two determinants of online intentional behaviours, it is concluded that trust in online forums has a wider role of influence on the online voluntary contribution.

However, contrary to what is suggested by H4a, findings show that trust in members does not impact on members' attitudinal intention. Jiang *et al.* (2002) argue that social recognition is a predictor of attitudinal intention in the context of knowledge sharing. Social recognition is found to be associated with interpersonal trust (McKnight and Chevery 2002). It is also likely that members who are popular within online forums are trusted more often because they are knowledgeable / competent. This understanding is embedded in the argument that knowledge availability is the reason why online forums exist (Wasco *et al.*, 2009). Against this, it is noted that ability based interpersonal trust is more difficult to measure than the benevolence / integrity based interpersonal trust (Levin *et al.*, 2003), and it is the case that ability has no significant prediction power on trust in members with the measurement models developed in this study. Thus, it is plausible that the unsupportive of H4a may be one consequence of the measurement validation, when trust in members and others antecedents of determinants are integrated to predict online intentional contribution behaviours.

Study one further explored the connection among the antecedents and how they affect each other. The findings showed that perception of a critical mass of membership within an online forum drives normative intention (H8). This builds on previous findings that have suggested that by perceiving a critical mass of membership, social pressure may lead individual members to want to be seen as 'normal' within a community (Cho 2011).

Previous studies had shown that interpersonal trust can promote a more general institutional trust (Luo, 2006; Zimmer *et al.* 2010). H5 is built on this by arguing that in the specific context of online forums, trust in other members influences trust in the online forum as an entity. The support of H5 suggests that

trust in members who are not behaving opportunistically favour the development of trust in online forums. Levin *et al.* (2003) argue that benevolence/integrity based interpersonal trust under the context of knowledge sharing is associated with the perceived knowledge available within an organisation. This argument is in consistence with the findings that the factor loadings of benevolence/integrity dimensionalities to trust in members are significant and convergent, while the knowledge available within online forums can reveal that online forums are able to and would like to allow members to access the digital public goods (ability and benevolence dimensionalities of online forums). The findings provide a possible explanation of the dimensionality of interpersonal trust influencing the institutional trust in the context of online forums.

The hypothesis of trust in members positively affecting perceived critical mass is not supported (H9a). However a more detailed examination of the data through mediation and moderation analyses revealed that the effects of trust in members on perceived critical mass are fully mediated by trust in online forums. This again highlights the role of trust in online forums in predicting the determinants of online intentional behaviours.

The support for H9b that trust in online forums affects the perception of critical mass, suggests that trust in online forums encourage members to voluntarily contribute knowledge, leading to mass numbers of usage of that forum. Similar results were found in previous studies (Granovetter 1973; Haythornthwaite 2002; Centola 2013). Trust in members is found to be a good predictor of trust in an online forum as an institution, and this drive intention to contribute knowledge.

Taken together, results show that the three identified antecedent effects on online intentional contribution behaviours in different magnitude levels, among which trust in online forums has more weighted power in the prediction of the online intentional contribution behaviours. Study one discusses the reasons for individuals continuously contributing and sharing knowledge in online forums, allowing the forum to remain a sustainable entity.

4.3.5.2 Theoretical implications

The motivation for study one is the absence of previous research which has taken an integrative view in trying to understand online knowledge sharing behaviours. Using a framework derived from TPB (Ajzen, 1991) and DTPB (Tylor and Todds, 1995), study one developed an integrative framework that investigates how contextual antecedents interact and influence drivers of intention to contribute knowledge in online forums.

The contribution of study one is threefold. Firstly, it has identified influencers of intention to online sharing. The perspective taken in study one is comprehensive and it considers various antecedents. Conversely, previous studies considered isolated or limited numbers of online voluntary contribution behaviours embedded in different theoretical approaches (e.g. Shen *et al.*, 2013). Study one argues that all of these factors should be taken into consideration when analysing intention to share knowledge online, and that taking them individually only provides a partial perspective.

Secondly, study one considers the role of contextual antecedents in influencing the determinants for intention to online sharing. Previous studies have mainly investigated the causal relationships between the determinants and the response variable by developing theoretical hypotheses, and examining those using SEM techniques (e.g. Chen 2007; Shen *et al.*, 2013). Examinations of the antecedents of the identified determinants have been little addressed previously. For instance, limited studies have sought to investigate the relationship between interpersonal trust and normative beliefs (Jeffries and Becker, 2008). Findings from study one can add knowledge about how higher levels of benevolence/integrity based trust in members can result in higher levels of perceived normative pressures on members. DTBP frames the understanding of the antecedents of factors that drive voluntary contribution online. Study one proposes the following antecedents for the determinants of intention to contribute knowledge within online forums: trust in online forums, trust in members and perceived critical mass. In particular, study one suggests that trust in online forums should be considered as a motivator for attitude and perceived behavioural control (as supported with H4b and H6); trust in members and perceived critical mass should be specified as

antecedents for subjective norms (supported by H7 and H8). Related research has previously found that trust in online communities can lead to a positive attitude to purchase online (Zimmer *et al.*, 2010) and that trust in the competences of online communities is positively associated with perceived behavioural control (Erden *et al.*, 2012). Given the fact that trust in online forums positively affects the two determinants of online intentional behaviours, it may be concluded that trust in online forums has a wider role of influencing online knowledge contribution.

Thirdly, study one analyses the impact that contextual factors may have among themselves. Again, this can be better examined within an integrative model, and has not been examined previously. With study one, the underlying relationships between the antecedents of determinants are further tested, and this helps to understand how these causal factors impact on the determinants for online contribution intention. To the knowledge of the author of this thesis, no previous research has sought to test the causal relationships between interpersonal trust and perceived critical mass. Study one has found that trust in members who would like to share knowledge with others can be a predictor of perceived size of contributors, but its effects are completely mediated by trust in online forums.

When addressing the way contextual antecedents act together to influence the determinants for online contribution behaviours, study one proposes that strong and weak ties co-exist, and they play a role in the continuance of online knowledge sharing. Although trust in online forums (involved with weak ties) has been found to have more power in predicting the determinants for online intentional behaviours, and it completely mediated the effects of perceived critical mass regressed on trust in members, the role of strong ties cannot be ignored, which is demonstrated by support for H5, i.e. trust in members (involved with strong ties) can lead to trust in online forums. That is, previous findings generated with simulations (i.e. Centola, 2013) that strong ties can impede the expansion of memberships are questioned with this empirical study. However, it is worth noting that the finding is in agreement with the argument by Granovetter (1973) that weak ties are more likely to occur if strong ties are present. Additionally, study one has further clarified that trust in members based on benevolence and integrity based are the

antecedents of trust in online forums, because the former is associated with the perceived knowledge available within a forum that can reveal the ability of that forum, in agreement with the findings by Levin *et al.* (2003).

4.3.5.3 Managerial considerations

For firms that host online forums in order to benefit from an open knowledge source, it is essential to pay attention to the role of users' perceptions of the 'health' of the atmosphere within the forum. This can comprise the vigour of knowledge sharing activities, the ability of technical support, the benevolence of others, and the presence of effective privacy protections. With these conditions in place, members of an online forum are more likely to be encouraged have a sense of belong to the forum. Where a forum is linked to a firm's commercial activities, this may facilitate maintaining the relationships with its customers. Secondly, these conditions can promote the expansion of knowledge contribution activities so that sustaining an online forum is plausible.

The importance of online institutional trust is revealed in the findings. It is also propose to firms that trust building among members can lead to the institutional brand building, hence further encourage a wider attitude to collectivism and collaboration. Although study one defines the scope excluding situations where knowledge is exchanged for direct financial reward, firms can nevertheless provide incentives for contribution. These can typically include the provision of opportunities for learning, and enhancing key contributors' reputation and status by publicly acknowledging their efforts (Hertel *et al.*, 2004). It is because the availability of knowledge is essential for online forums (Wasco *et al.*, 2009), firms can select members for promotion based on their numbers and quality of knowledge contributions.

Researchers argue that some moderation and leadership support are imperative for online communities (e.g. Erden *et al.*, 2012). The results show that the subjective norm that is influenced by trust in members and perceived critical mass is a significant predictor of members' intention to contribute knowledge online. Hence, it is argued that some moderation activities should be associated with achievement of this, for example, by facilitating IT support which can allow members to establish relationships among

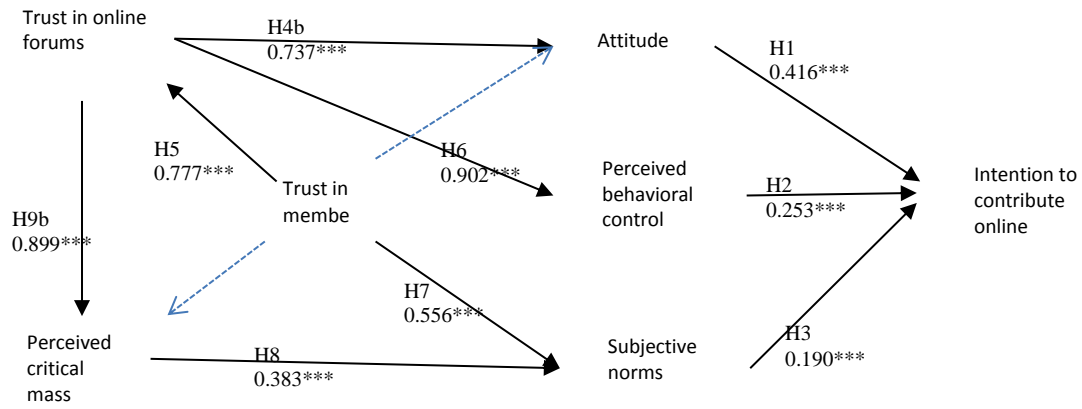
themselves independently and easily, and allowing members to feel psychologically safe and to be able to freely express their ideas. Firms are urged to be benevolent and integrative, because this is the plausible way to facilitate the creation of a “healthy” atmosphere within an online forum.

4.3.5.4 Limitations and further research

There are several limitations in the study. Firstly, it is subject to the limitations of a structured, quantitative approach to data collection which can be associated with issues of interpretation of the questionnaire by respondents (Hsu *et al.*, 2007). Although words used in the survey have been carefully selected, it is possible that some respondents understand the questions differently to what was intended and this could impact on interpretations of the final results. However, against this, the sample size (900) is sufficient for SEM analyses according to the 10:1 rule-of-thumb (Nunnally, 1967), and can help to reduce response bias. Secondly, the dynamic aspects of trust and critical mass concepts are investigated through an online survey, which only explains explanatory characteristics of these antecedents. Thirdly, this study has adopted a deductive, quantitative approach to investigating the issues of interest, and a deeper understanding of the relationship between phenomena may be derived from further complementary inductive, qualitative approaches.

A good area of future research could examine how trust is evaluated online and how network structures impact on sustaining online communities (study two and study three). A longitudinal experimental design – although potentially difficult to develop – may enhance the predictive power of the model developed in study one.

Figure 20: Final result



*** $p < 0.001$; → supported paths; --> unsupported paths

Chapter 5 Results and discussions: Study two - the development of online trust

5.1 Development of themes

The form of analysis used in the inductive phase is based on the principles proposed by Miles and Huberman (1994), and Strauss and Corbin (1998), and is iterative in nature. Data collection and analysis are consciously combined, and the initial data analysis is used to guide ongoing data collection and coding. Reviews are coded and analysed using NVivo software. This allows the researcher to identify associations between themes of comments, and to supplement this with contextual data introduced by the researcher.

To develop initial themes, free nodes are coded using themes derived from the first 10 randomly selected reviews. Free nodes that shared common underlying ideas are merged into tree nodes. The list of tree nodes gradually expands as more reviews are analysed and represented in the list of categories and themes that emerged from the coding. Following the recommendation of Gibbs (2002), coded nodes are arranged into a node tree. Tree nodes are located at the top, representing emerged themes, beneath which are lists of child nodes and sub-tree nodes (Gibbs, 2002).

Based on the similarity of meaning, four themes emerge from the process of initial free coding and the subsequent development of tree nodes: ability, benevolence, integrity and predictability. The comments relating to the emerged theme are rated on a scale from “1” to “5”, with the code “1” representing a very negative comment, and the code “5” referring a very positive comment.

Ability is understood as the perceived competence of the company in meeting the promises it makes to customers. The emerged theme of ability includes items identified in the literature of perceived ease of use, perceived usefulness (Davis *et al.*, 1989; Koufaris and Hampton-Sosa, 2004), and perceived quality (Delone and McLean, 2003). The followings are typical of comments relating to this dimension:

- “Skype’s service is undeniably brilliant, simple, convenient and cost-effective. its Ultra Clear Sound, Free, Easy To Use and More !!” (Coded 5).
- “Occasionally, the voices will sound like they’re underwater, or go crackly, but this doesn’t happen too frequently” (Coded 3).
- “I was VERY scared. I didn’t know what it was then I figured it was the Skype ... BAD! I would advise disabling the sound effects on start-up and shut down of the program” (Coded 1).

The benevolence dimension of trust is understood as the willingness to help users solve problems and obtain maximum benefits from the service. The following are typical of comments coded as representing the theme of benevolence:

- “This feature is unique in itself, you cannot get it anywhere else than skype” (Coded 5).
- “But for me the issue which makes this product a firm NO! is the company’s complete lack of willingness to deal with the security issues, and their actions towards worsening this” (Coded 1).

The third theme of integrity is understood as a general willingness of the organisation to keep promises and to act in an ethical manner. The following are typical of comments coded as representing the theme of integrity:

- “This free software does exactly as the title suggests” (Coded 5).
- “I had my account blocked for no reason. I can’t take any chances on Skype. I’d rather pay more to know that my account is secure. If they’re going to block me, at least wait till I finish my credits first. I had been a long-time customer of Skype since 2006, but I will not do business with them again” (Coded 1).
- “When I first heard about Skype I thought it sounded too good to be true, and it did not take me long to find it was all too much hassle” (Coded 2).

The fourth emerged theme is predictability, which is understood here, as a general belief that the company will act in a predictable and consistent manner in its dealings with customers:

- “The software is updated every so often, with each version ironing out any glitches with the previous version” (Coded 4).
- “Many worry about others listening in on calls and to add to this worry, Skype will either admit or deny this claim” (Coded 1).
- “There is a POSSIBILITY that it will be updated, but it seems to be a very slim one as I saw no evidence of it in the time I used the application” (Coded 2).
- “Personally I do not trust the Skype system for sensitive information” (Coded 1).

In addition to the score given for each of these dimensions of trust, a score is given for overall trust. This is the overall assessment of the extent to which the comments expressed by the contributor indicated whether or not they trusted Skype. Table 41 provides a summary of the number of times each emerged dimension of trust was mentioned in the 352 reviews studied. In the summation, comments which mentioned more than one dimension of trust are counted for each dimension.

Table 40: Numbers of comments by emerged themes, 2005--2010

	Ability	Benevolence	Integrity	Predictability	Total coded items referring to the dimension of trust during year
2005	80	60	33	8	181
2006	26	23	10	6	65
2007	53	49	20	3	125
2008	98	65	27	9	199
2009	182	112	58	15	367
2010	95	48	40	11	194
Total	534	357	188	52	1131

It is noted in the methodology section that a number of distinctive events could have contributed to changes in trust in Skype during the period from 2005 to 2010. A content analysis of news media items relating to Skype for the period 2005 to 2010 is undertaken and items analysed to assess whether widely reported stories corresponded to turning points in observed trust. Three sources of news are analysed: BBC News (www.bbc.co.uk); CNN News (www.cnn.com); and Cnet News (<http://news.cnet.com/>). These are felt to represent a good balance between mainstream press and specialist technical news media, and provide a good geographical spread of target audiences. Further analysis is undertaken of scores for each of the emerged dimensions of trust, in the section 5.2.

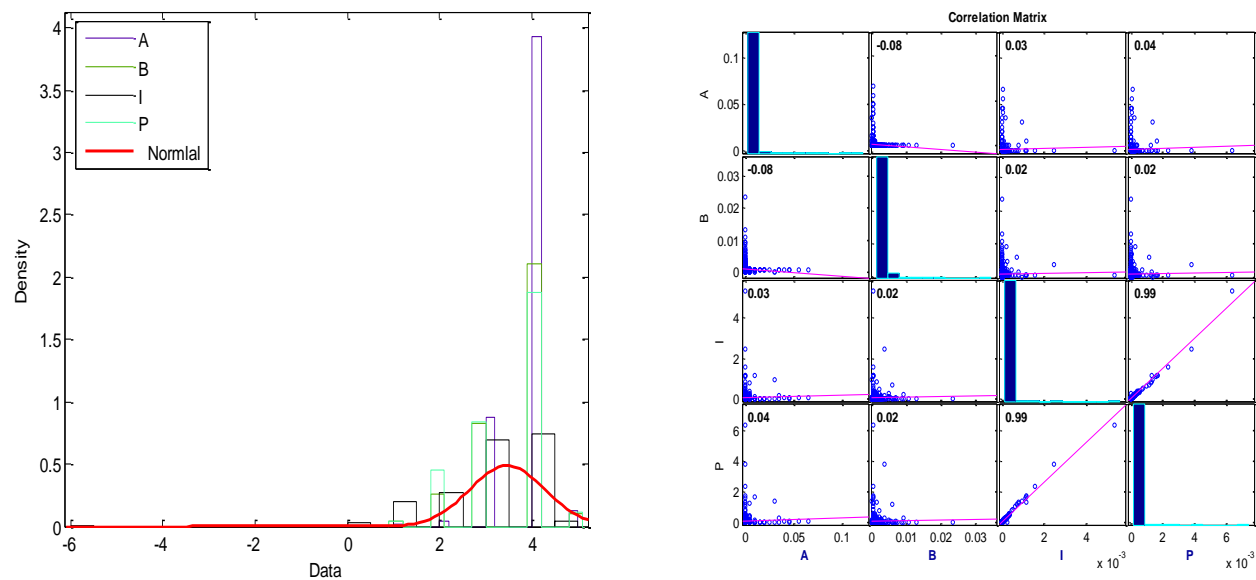
5.2 Exploratory analysis

Four themes emerge from the inductive phase. Results from coding are further analysed in order to understand whether the pattern of the emerged themes is not random, and the emerged dimensionalities of online trust are distinctive from each other. This infers a deductive logic that is followed.

The distribution of the sample data is significantly different from the normal distribution. Moreover, the Pearson's correlation tests show that the predictor variables, "integrity" and "predictability", are linearly correlated, with p-values greater than the boundary of 0.05; the predict variables, "ability" and "benevolence", are also linearly correlated. However, Pearson's correlation coefficients equal to -0.08, suggesting an opposite relationship.

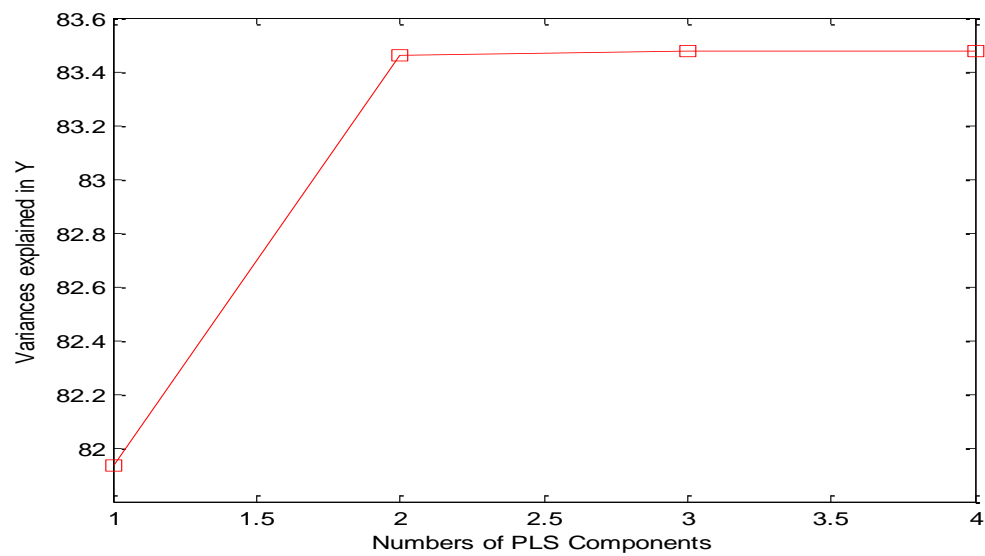
Partial least squared PCA is firstly conducted in order to address the issue of multicollinearity among predictors. This method helps to select the numbers of predictors by measuring the covariance between the responses values regressed on the predictors.

Figure 21: Data distribution and correlation



A: ability; B: benevolence; I: integrity; P: predictability

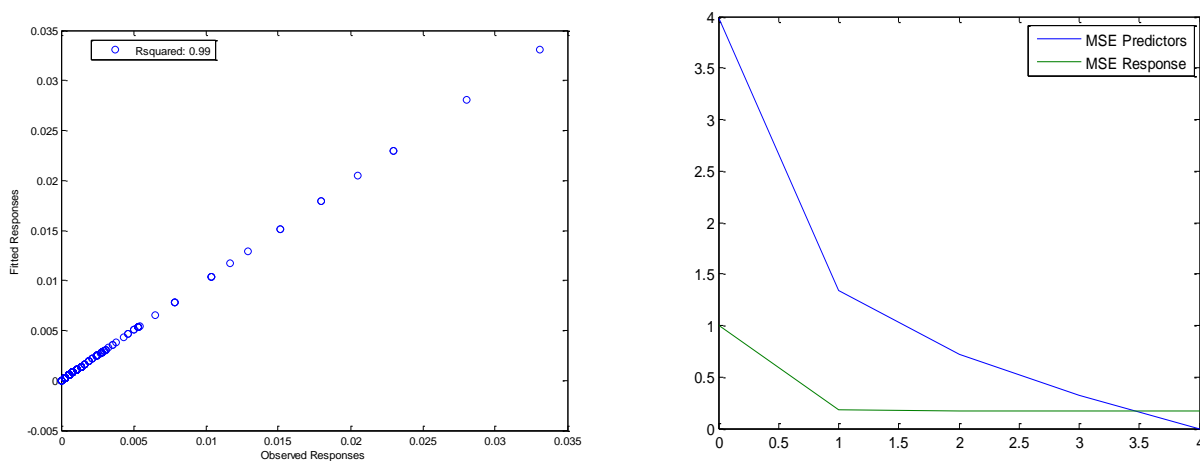
Figure 22: Components extracted



The predictor “ability” contributes nearly 81% of variance explained in the dependent variable (online trust), and the rest of the predict variables “benevolence”, “integrity” and “predictability” have the similar

contributions to the variance explained. This may be a result of the unequal numbers of codes for the emerged themes. The exploratory data analysis by PLS regression suggest that the numbers of predictors can be reduced into 2, “ability” and “benevolence”. With the numbers of predictors reduced to 2, the correlation between the observed and fitted responses is almost linear, with R-squared (0.99). The mean squared error in responses is not changed significantly by excluding or including the predictor “integrity” and “predictability”. The results show that “ability” and “benevolence” have more prediction power for the response variable “online trust”.

Figure 23: Fitted responses and MSE



However, the variances explained in the response variable by the four predictors are only up to around 84%, suggesting that the data may be better presented in a higher dimensional space. The next stage is to find a classification method that can study the relationships between the predictors and response variable in the high dimensional space. In this study, SVM is used for classification.

5.3 Dimensionality of online institutional trust

Analysis proceeded from the inductive to the deductive stage by seeking to establish whether the trend pattern of the emerged themes was non-random to establish whether the emerged themes were distinct components regressing to online institutional trust.

Three SVM models are trained and embedded in the different kernel function, i.e. linear, polynomial and Gaussian. The samples are randomly divided into 10 folders. The trained models are tested on 9 folders, and the 10th folder for the testing validation. The average errors over all testing are computed and are used as the major criteria for the model selection. The main reason for using SVM models are that they are more flexible in dealing with the issue of multicollinearity by using the technique of projection. That is, the raw data set is seen as five vectors (4 predictors and 1 responsible variable) and analyses are undertaken with their projections.

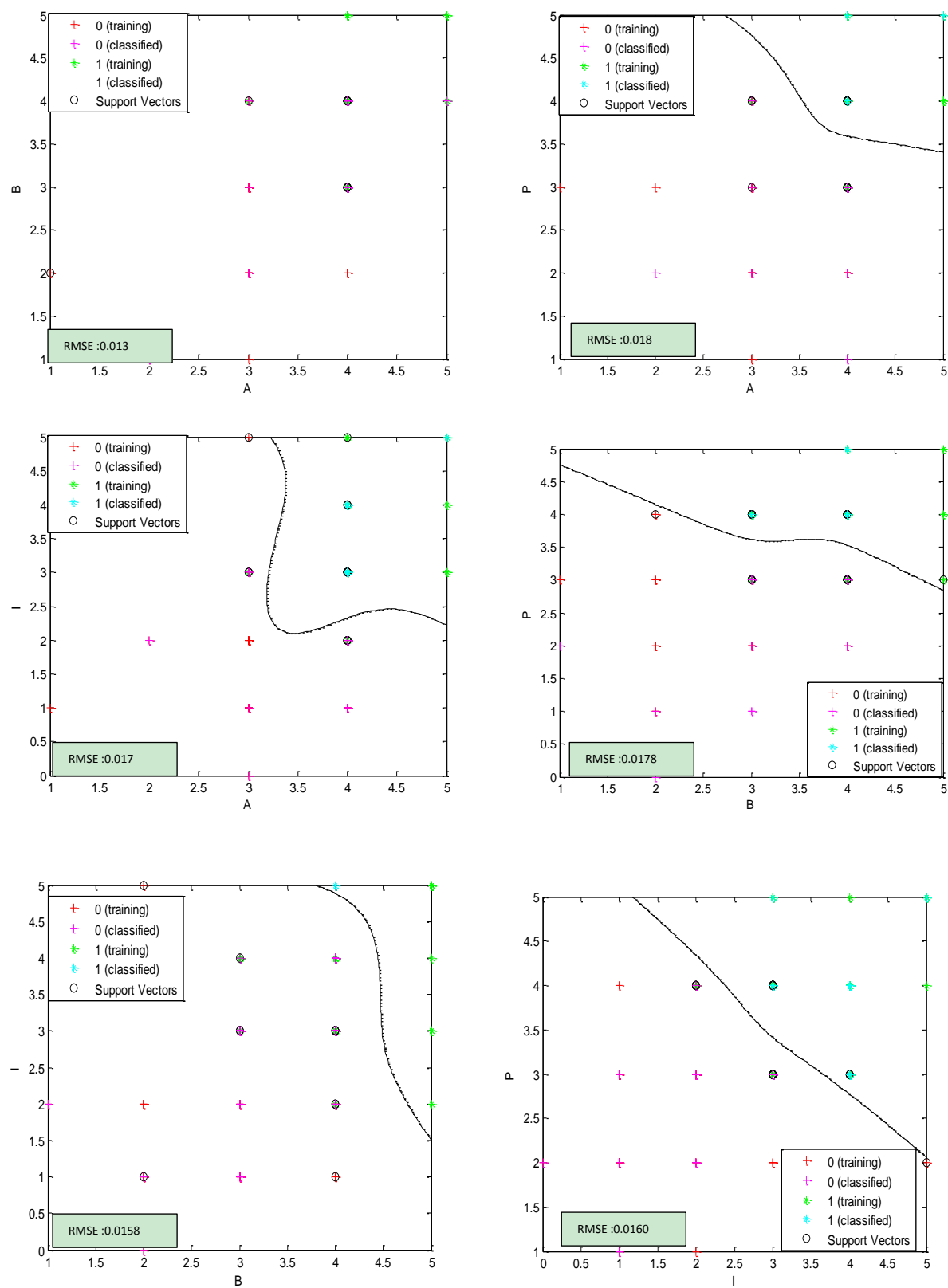
Table 41: *Comparing the kernel function*

Kernel Function	Average errors	Numbers of support vectors	Bias
Linear	0.1073	43	-0.2489
Polynomial	0.1030	47	0.0710
RBF(Gaussian)	0.1459	91	-0.3278

The polynomial-SVM has a better performance, with the lowest average errors at 0.1030. This suggests that around 10% of data has weak predictive performance. One possible explanation to the data loss is that the raw data is highly skewed, but deleting a fraction of outliers (for example 1%) can lead to the higher value of average loss.

The following plots show how much similarity existed between the pair predictors in measuring the response variable. Online institutional trust is divided into binary groups, “weak” and “strong”. The support vectors are those on the decision boundary for classification. It shows that the four predictors can be distinguished from each other. In particular, “integrity” and “predictability” are less biased in predicting either “weak” or “strong” online institutional trust. However, “ability” and “benevolence” are confounded in measuring the “weak” online institutional trust. This may be inconsistent with the results from PLS that they can explain around 85% of variances explained in the online institutional trust. Analysis suggested that the emerged themes of ability, benevolence, integrity and predictability are distinct components of online institutional trust.

Figure 24: Paired predictors



A: ability; B: benevolence; I: integrity; P: predictability

The posterior probability against the influence of the individual predictor is computed. Appendix 3 predicts the distribution of the joint probability that paired predictors contribute to the scores of general trust. The posterior probability ranges from 0 to 1, with 1 representing a situation that the strongest trust is more likely to occur, and 0 referring to the weakest. For each combination, the 3D plots (on the top) show that the joint influences on trust are not linear. The contour plots (on the bottom of the 3D plots) illustrate how the scores of the predictors are distributed, and their geometric positions on the 3D plots are distinguished by the same colours.

Results showed that when “ability” and “benevolence” have the coding levels at 4 or 5, it is almost sure that they can lead to “strong” institutional trust. “Integrity” and “predictability” have fewer effects on the “strong” online institutional trust, because their coding levels can range from 1 to 5. However, the “weak” levels of the responsible variable are more likely to be measured with “benevolence”, “integrity” and “predictability” (for example, coding 1 -2 for “benevolence” lead to ‘weak’ trust levels, but “ability” can have the coding from 1 to 4). In addition, it is almost sure that the low levels of “integrity” and “predictability” can lead to ‘weak’ online institutional online trust.

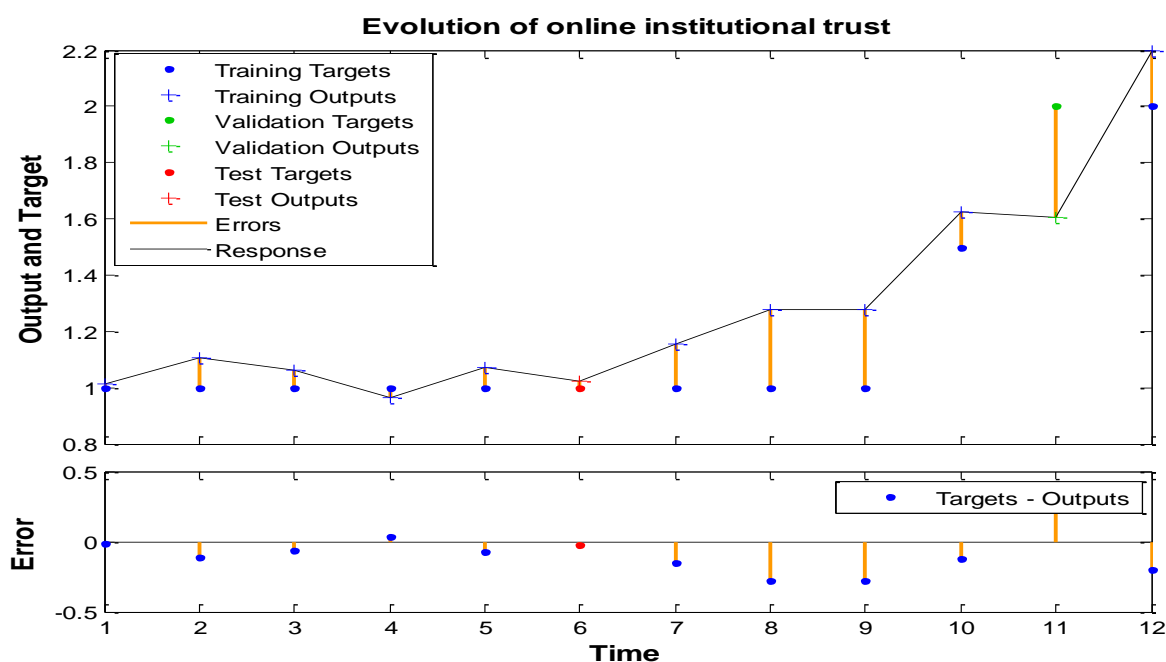
The result is in agreement with the arguments by Domikia (2010) that the initial trust building is more likely to be associated with the cognitive factors, such as “ability”; while the undermining of trust is more likely to be influenced by the affective area of brain, such as “benevolence”.

5.4 Evolution of online trust

The above results show that “ability”, “benevolence”, “integrity” and “predictability” are distinguishable components of online institutional trust. The scores for the evolution of online institutional trust are then analysed. Initial inspection of the scores for overall trust indicated a slowly rising pattern for the year 2006, followed by a more rapidly falling pattern in 2007 to 2008. The overall trust score raises again in the year 2009 and falls in 2010. This indicated a discontinuous trend in trust during the period of study, but most notably, the points of discontinuity occurred at different points for each of the emerged dimensions of trust.

The autocorrelation outputs show that there is only one error correlation extending over the 95% confidential levels, and it occurs at the zero lag. This suggests that the prediction errors are uncorrelated with each other, and the feed forward model is adequate, recalling that the model is $y_t = f(y_{t-1}, \dots, y_{t-d})$.

Figure 25: Evolution of overall trust in online forums



The recurrent neural network model is not over fitted, because the test curve does not increase before the validation curve, and the test curve has the similar trends with the one for training. The regression plots show that the relationships between the output and the target are linear, with R closing to 1. In other words, their relationships are more likely to be correlated but not happened in random.

Figure 26: Goodness plot: evolution of overall online institutional trust

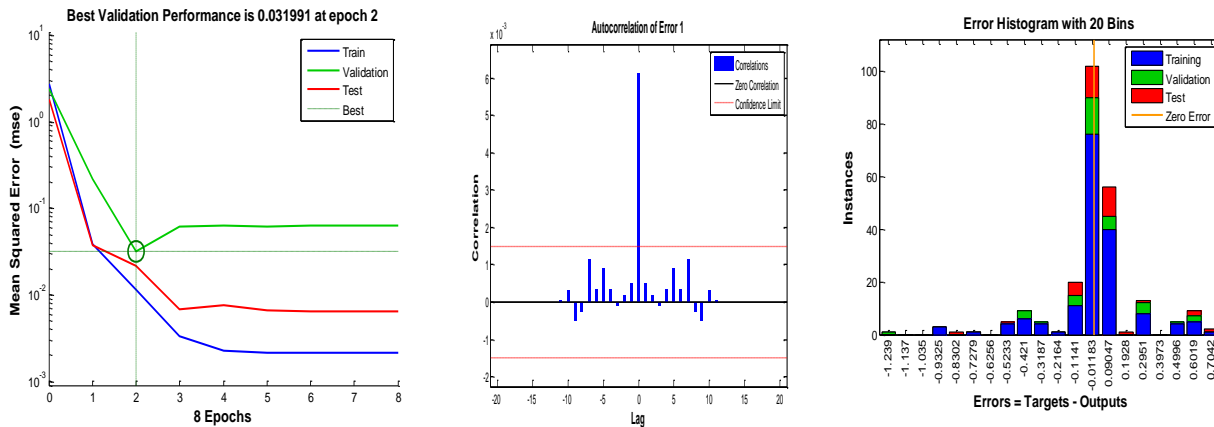
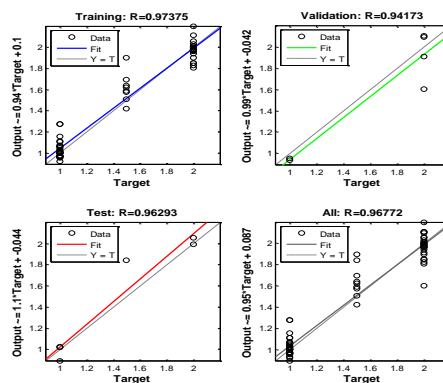


Figure 27: Linearity plot of overall trust in online forums



Three widely reported news stories were noted that could have had an effect on trust in Skype during the studying period. Firstly, there was extensive reporting of a major and prolonged systems failure that occurred in August 2007 and which led many reporters to seriously question for the first time the trustworthiness of Skype to deliver a consistent and reliable service on mobile phone. A second widely reported story was the takeover of Skype by e-bay in 2006, accompanied by predominantly negative comments of a big corporation stifling the entrepreneurial zeal of Skype's Baltic founders. The third story happened in 2010 when Microsoft took over Skype. Users worried that they might have less control of their private information. However, neither of these events was closely related to turning points of overall trust which occurred in subsequent years.

The results for the predictions show a climbing trend for the coming 6 years after the study period. This is in consistent with the stories by news resources, that Microsoft has advanced its ability in response to users' requirements, such as using universal translators for video and messages exchanged within users, and developing a mobile platform and partnerships with Facebook. The Skype number price increased 1600% in the year 2015. It shows trust-building encompasses a number of processes which typically take time to develop.

The same model equation is applied to each component of online institutional trust in order to understand how they change over time. The goodness of model fits (e.g. Error histogram) suggests the adequate model fit. Results show that “ability” plays an important role in the initial trust building (e.g. first peak at year 2006 when the overall trust scores increase for the first time).

Figure 28: Evolution of ability

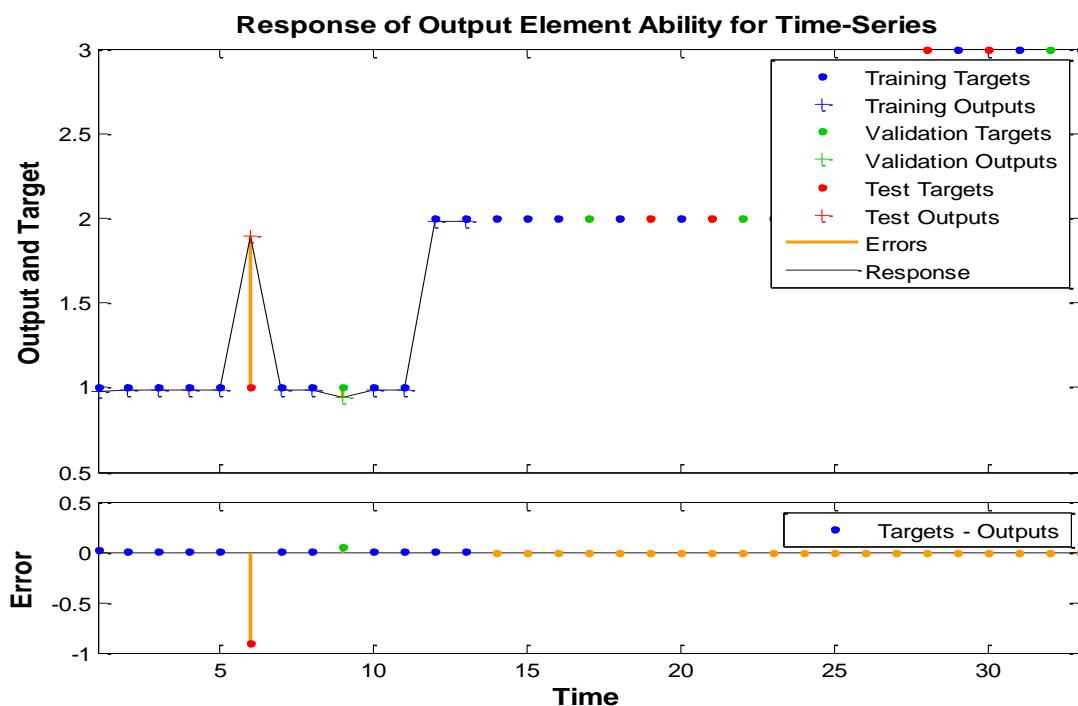
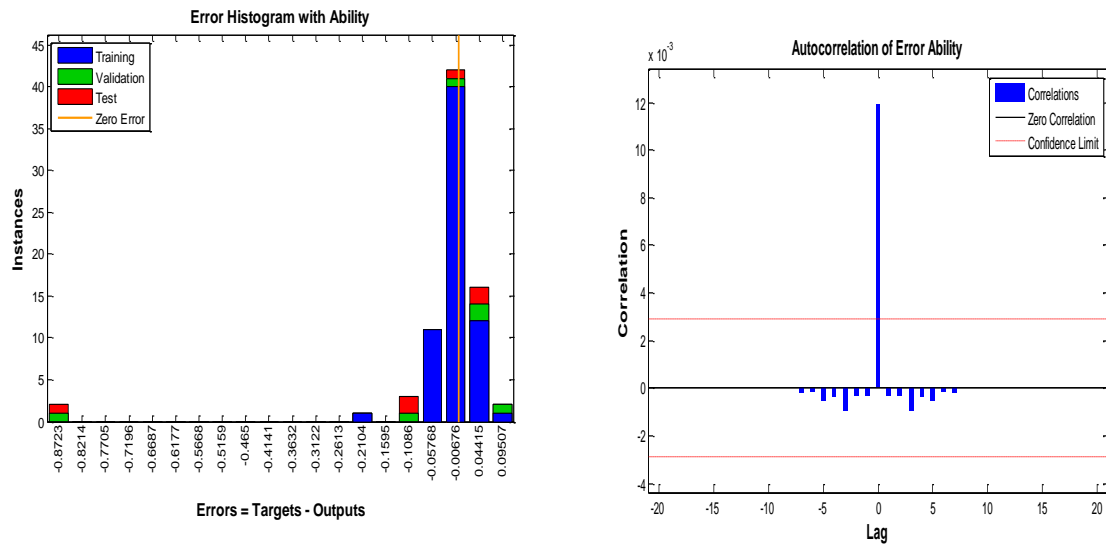


Figure 29: Goodness plot: evolution of ability



“Benevolence” and “integrity” are associated with the decline of the overall trust scores (i.e. a significant decline of the level of “benevolence” in 2010, the year when overall trust scores are the lowest during the studying time). Moreover, “benevolence” is important for the overall trust building in a later stage (i.e. increase with the overall trust scores in the forecasting period).

Figure 30: Evolution of benevolence

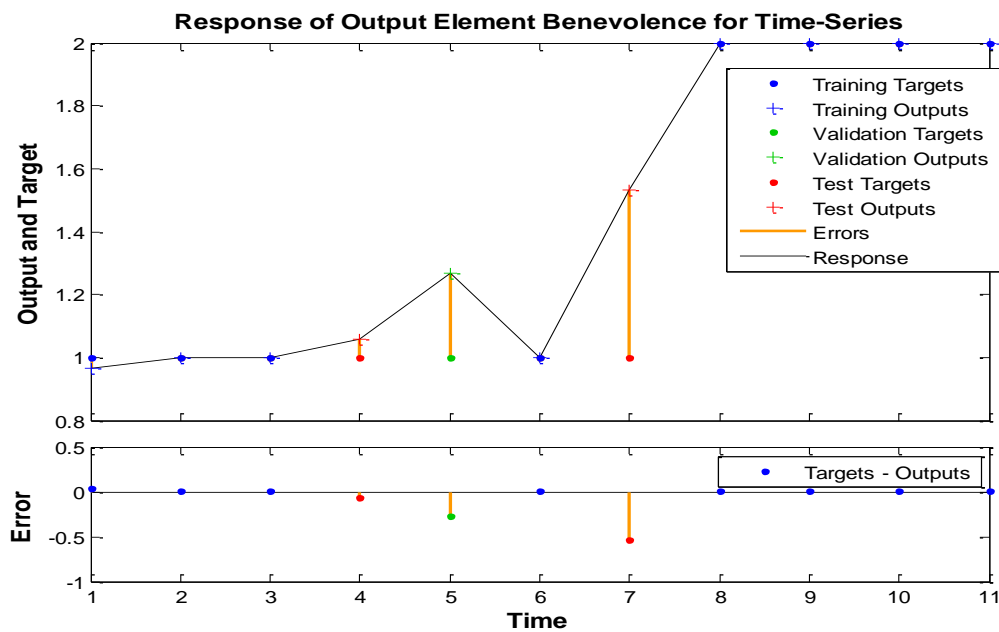
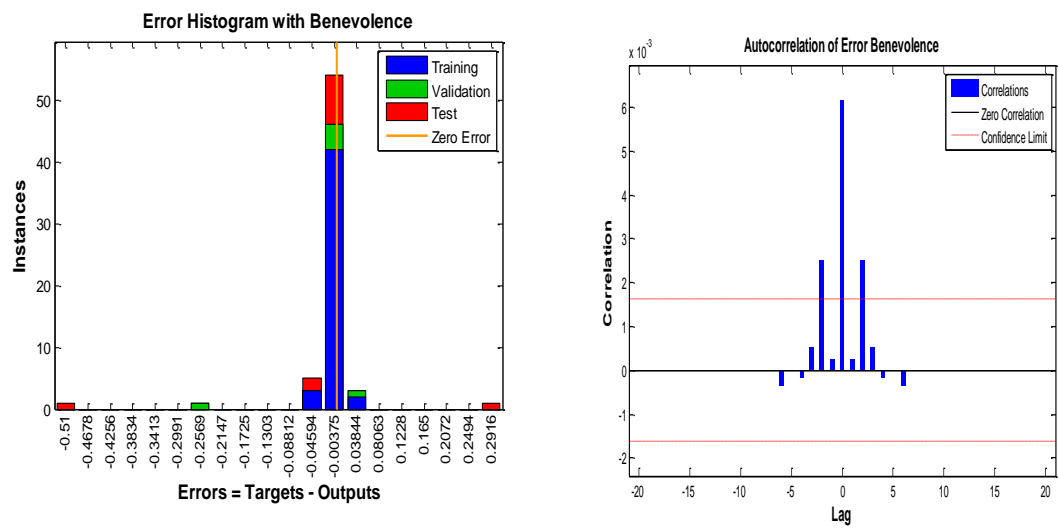


Figure 31: Goodness plot: evolution of benevolence



“Predictability” and “integrity” have fewer effects on the initial trust building (i.e. all flat during the first period). However, the increase of “predictability” is associated with an overall climbing trust scores (i.e. the level of “predictability” increases in the later stage).

Figure 32: Evolution of Integrity

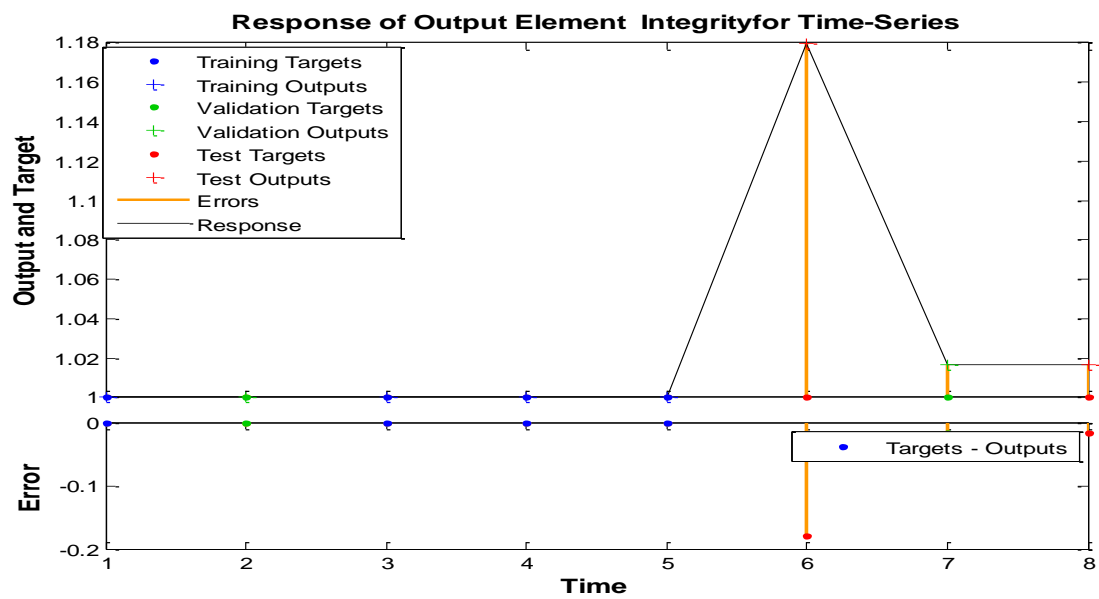


Figure 33: Goodness plot: evolution of integrity

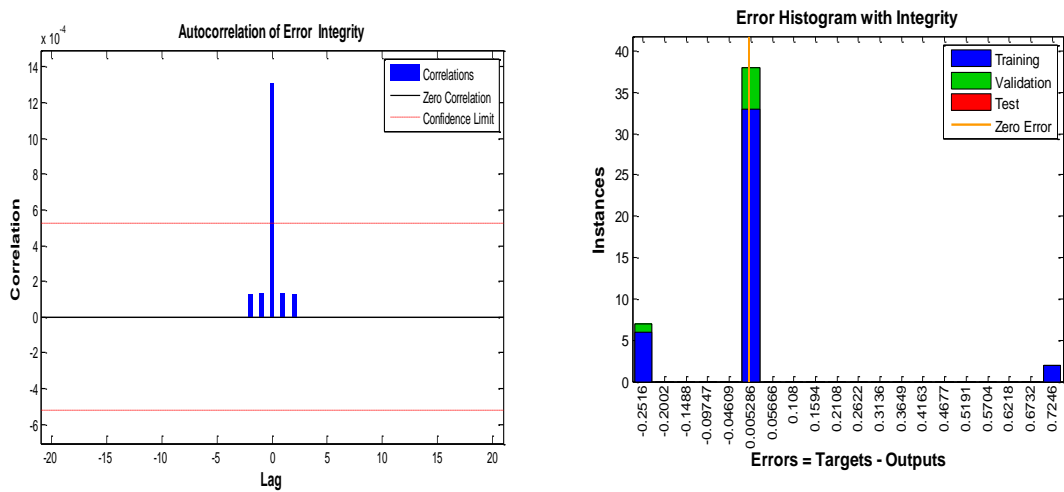


Figure 34: Evolution of predictability

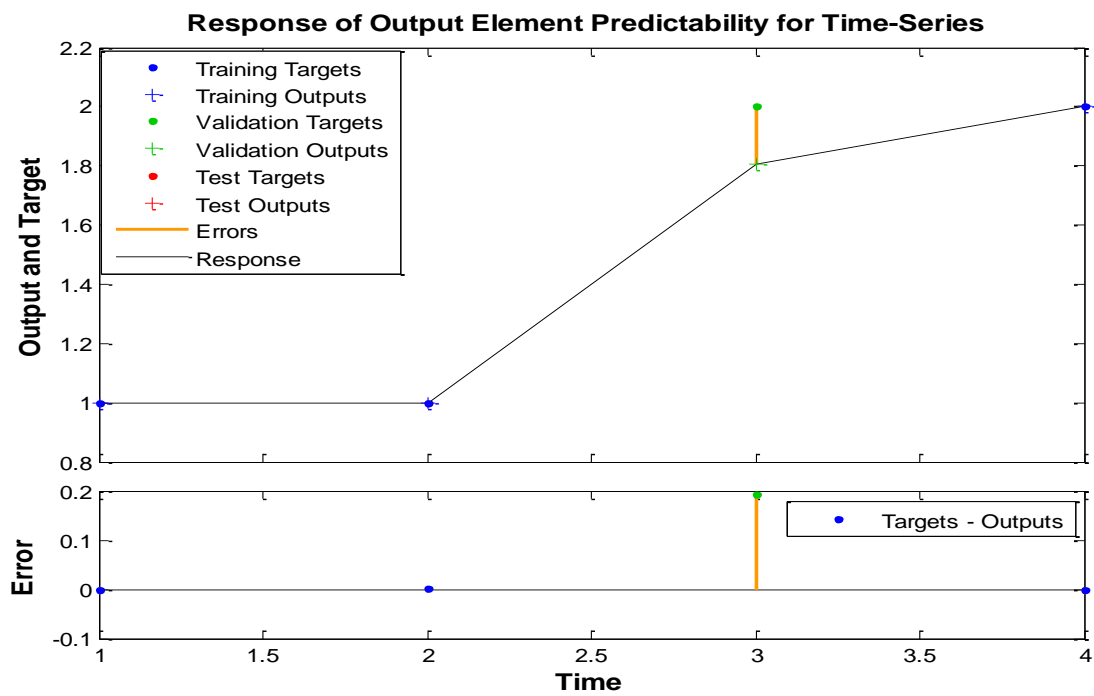
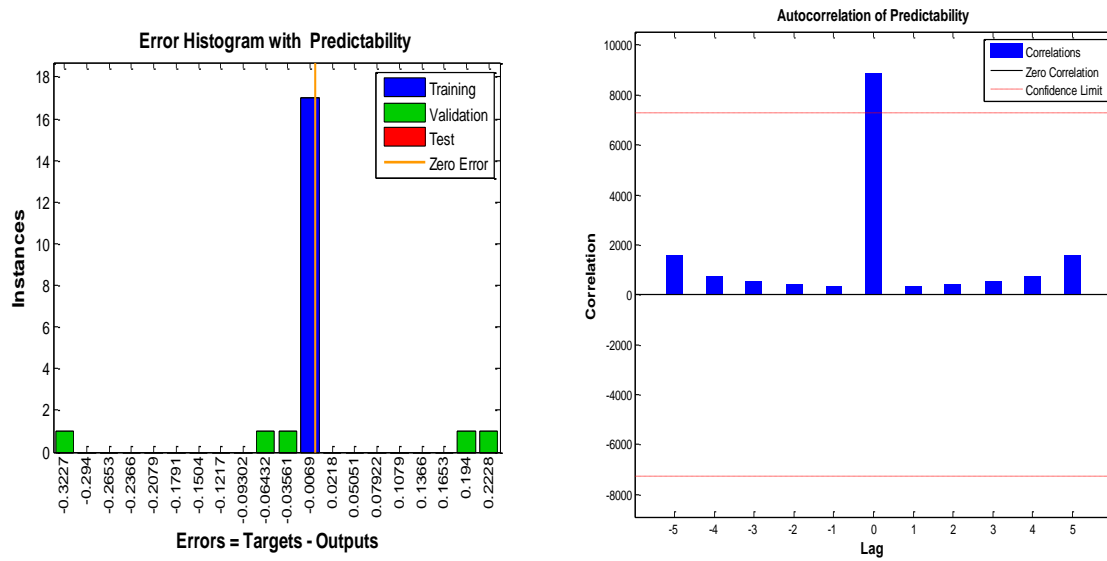


Figure 35: Goodness plot: evolution of predictability



Results from the neural network analyses are similar to that from SVM. However, the neural network approach is more sensitive to the issue of sample size and missing data. It is because the numbers of coding for ‘integrity’ and ‘predictability’ are smaller than that for ‘ability’ and ‘benevolence’, the results are plotted, remaining in the study period for these two components of online institutional trust, and have less predictive power. However, the neural network approach can be seen as complementary to SVM, because it is useful in understanding the evolution of each component.

5.5 Conclusions

Results generated from study one show the importance of the online institutional trust, but gaps exist in knowledge of how it evolves over a period of time. This study has made a contribution to theory and to marketing management practice by analysing users’ perceptions of the components of trust in a brand and noting how these components change over time.

The findings are consistent with the technology acceptance model (TAM) (Davis *et al.*, 1989), which advocates that perceived ease of use and perceived usefulness are the cognitive factors that facilitate users

to employ a new technology, and which play a less important role once the technology is accepted. “Ability” was seen to play a key role in initial trust development in a start-up brand. In the second stage, other affective factors such as the fear of exposing too much personal information, may determine users’ intention to use, or to continue to use, a new technology. Scores for the “ability” dimension of trust – which primarily involves cognitive evaluations – peaked earlier than the “benevolence” dimension, which involves more affective evaluations.

The findings provide additional support to the view that trust and distrust are distinct constructs, and that trust-building is associated with the brain’s cognition areas, while the undermining of trust is more strongly linked with the brain’s procession of emotions, particularly relating to deception, fear of loss, and unethical behaviour (Dimoka, 2010). The majority of previous studies have treated online trust as a unidimensional construct, and further research would be useful to explore the notion of online trust and distrust as two separate constructs, with distinct sets of antecedents and consequences. The cross-sectional approach using retrospective recall that has dominated previous studies may have impeded the identification of trust and distrust as distinct constructs, and it is a contribution of the methodology of this study which has suggested that distrust is a distinct construct. Further studies should be undertaken in different contexts to explore this suggestion further.

From the perspective of theory development, this study has added to the literature on the dimensions of online trust, and made a contribution by identifying how these dimensions change in salience over time. Previous studies into the dimensions of trust have not generally explored the stability of the dimensions over time, and this study has shown significant differences in their composition over time.

Methodologically, this study has made a contribution by using comments that were reported contemporaneously during a six-year period, thereby alleviating the problems of retrospective recall of trust, which have limited many previous studies of trust development, including the majority of the investigations of trust identified (e.g. Swan *et al.*, 1999). Qualitative techniques were used inductively to

build a theory that was then examined using machine learning techniques, and strong evidence was found of variation in levels and dimensions of trust in an institution over time.

From the perspective of practice development, the findings of this study inform online organisations' strategies for developing trust in a brand over time, particularly by identifying the dimensions of trust which need developing and reinforcing at different points during the brand's lifecycle. Promotional material and customer support in the early stages of a brand should seek to build trust in ways which are assessed cognitively, for example, by ensuring that the product performs to the standard expected by customers. At later stages, marketing strategy should place greater emphasis on customers' affective evaluations by providing evidence of an organisation's benevolence, for example, by demonstrating an empathetic approach to resolving problems when they occur.

The major limitations in the method are associated with the subjective coding, although coding is compared between 2 internal raters. Moreover, it is possible that the observed decline in trust may have been confounded by other cyclical processes. For example, the new products may typically be associated with excitement at the novelty of the product, and users may be prepared to forgive failures amidst the excitement of the novelty. Greater familiarity with the product may bring about greater critical awareness and willingness to publicly and privately complain about the shortcomings of the product. This possible lifecycle may be linked to trust, or it may act as an independent construct. However, these events did not by themselves appear to have a direct relationship with the turning points in any of the dimensions of trust.

The findings of study two are in agreement with those from study one. In study one, "ability" and "benevolence" are identified as two valid measurements of online institutional trust. In study two, besides to the above mentioned components, "integrity" and "predictability" are found significantly contributing to the overall online institutional trust scores. However, "integrity" and "predictability" have less power in predicting the evolution of online institutional trust. It is noted that the objectives of study one and study two are slightly different. Online institutional trust is identified as a key antecedent of online

knowledge contribution behaviours in study one. Study two has sought to learn the concept of online institutional trust in a dynamic view. Results from study two further found out the differential rates of change in the dimensions of trust, which peaked at different times, is an interesting phenomenon that should be validated with further research.

Chapter 6: Results and discussions of study three: the role of critical mass members in sustaining online forums

6.1 Introduction

Embedded in the results generated from study one, perceived critical mass is an important antecedent to online knowledge sharing. The concept of perceived critical mass in study one is measured through an online survey, with items adopted from previous literature and embedded in the ideas proposed by Oliver and Marwell (1988) who argue that a small group of members can evoke the mass collectives.

This concept is originally borrowed from the theory of self-organised criticality (Bak *et al.*, 1987) in the field of complex systems. The latter seeks to reveal how non-linear interactions between nodes (actors in social science) in particular the hubs (critical mass members in social science) can bring a phase transaction (such as mass collectivises in social science) in the circumstance of asymmetric information and local movements without central control. This context is similar to the background of contribution behaviours within online forums, as the online contributors are volunteers (without central control and asymmetric information), and the communications often occur between small groups of members (local movement) rather than everybody participating in the same discussion. Indeed, online forums can be conceptualised as complex network graphs (a branch of complex systems) consisting of nodes/vertices (human actors) and edges (social relationships), and imply structures marked by the emergence of hubs (critical mass members) (Dorogovtsev and Mendesy, 2002; Albert and Barabási, 2002; Newman, 2003).

Study three is therefore an extension of analyses of results from study one which takes the approach of network science, inferring a different methodology from structural equation modelling and webnography. Results from study one support two essential dimensions that form the concept of perceived critical mass, i.e. in addition to the perceived association with others members, the perceived density of communications that involves the size of online forums are supported with statistical significance. In the language of network science, the size of networks matters in the emergence of phase transition, and that some nodes play the role of “bridge” between interactions of nodes within networks. In this study,

network languages are employed with nodes referring to members, hubs/giant component(s) indicating the critical mass members, degrees meaning the linkages associated to members, and networks representing online forums.

This study is organised as follows: the structure of networks is explored and described after the description of data which is provided by Stanford University and publicly available online. The network structural analysis uses the computer language Python with respect to the suggestions by Clauset *et al.* (2009). The role of hubs/critical mass members is evaluated within the network characterised by scale-free of degree distribution through attending and random attacks to the network. Finally, the evolution of a network is investigated.

6.2 Descriptive data

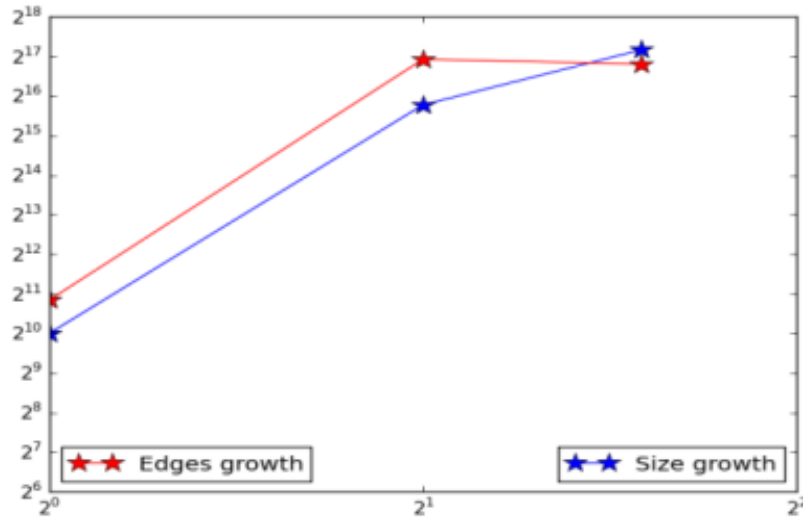
The online forum “Stack Overflow” is a website where members frequently seek help from others by posting questions. Answering questions posted by others is typical of voluntary contribution behaviours (Wasco *et al.*, 2009). Stack Overflow provides user generated data including IDs for both question and answer owners during a period of 71 months. The top questions asked and answered are about Java language, which represents a direct network consisting 146874 nodes and 333608 edges (<http://snap.stanford.edu/proj/snap-icwsm/SNAP-ICWSM14.pdf>).

With this database, an online “contribution” network is created with the help of the data visualisation software Gephi. Gephi is used to transfer the weighed direct network to an unweighted and indirect network, which produces a new data set with 147190 nodes and 149289 edges. There is an increase in the numbers of nodes with the new dataset as Gephi will automatically add missing nodes into a graph for analyses. There is a decrease of numbers of edges as an unweighted indirect network counts one time edges between paired nodes. The original data set contains a regular cycle. A weighed network can describe intense communication and is therefore often studied in relationship networks such as Facebook. Online forums represent interest oriented networks, where contribution of knowledge is not likely to be embedded in the assumption that contributors know the question owners.

In addition, online “contribution” network is undirected in the sense that contribution behaviours are correlated with topics that connect question owners and contributors. Direct networks involve in-degree (the numbers of edges toward question posters) and out-degree (numbers of edges from question posters to answers). Thus, the in-degree reflects the contributing activities of members, while the out-degree can suggest the interests of contributors. If the topic is interesting to the contributor; one may be more likely to participate in the online discussions. According to critical mass theory (Olivier *et al.*, 1985), contribution is also influenced by the production function of digital public goods (knowledge available to everyone), which reveals the unbalanced relationships between output (responses to questions) and input (questions). It is because contributors have more resources to contribute (presumably being more knowledgeable), that the digital public goods (knowledge available from Stack Overflow) are more likely to be created by contributors who have more resources (knowledgeable members’ contributions). In this sense, out-degree should be correlated with in-degree because out-degree can reflect the reason why experts contribute in Stack Overflow. In other words, the “contribution” Stack Overflow network is an unweighted and indirect network created from the same direct network by ignoring the direction of edges.

The sum of degrees is 149289×2 , with the highest degree at 2275. Less than half (34.24%) of members who contribute 100% of network effects have degrees different from zero. 0.417% members with connections at least equal to 42 contribute 30% of network connectivity effects. The membership growth exceeds the edges’ growth, which tends to slow down, suggesting that many members have few or zero connections. That is, the expansion of membership is not proportionally growing with the contribution behaviours. This provides an initial observation that information spreading is made by a small proportions of members.

Figure 36: Descriptive data: an overview of edges and size growth



6.3 Exploration of “contribution” network structure

6.3.1 Online forum Stack Overflow is scale-free

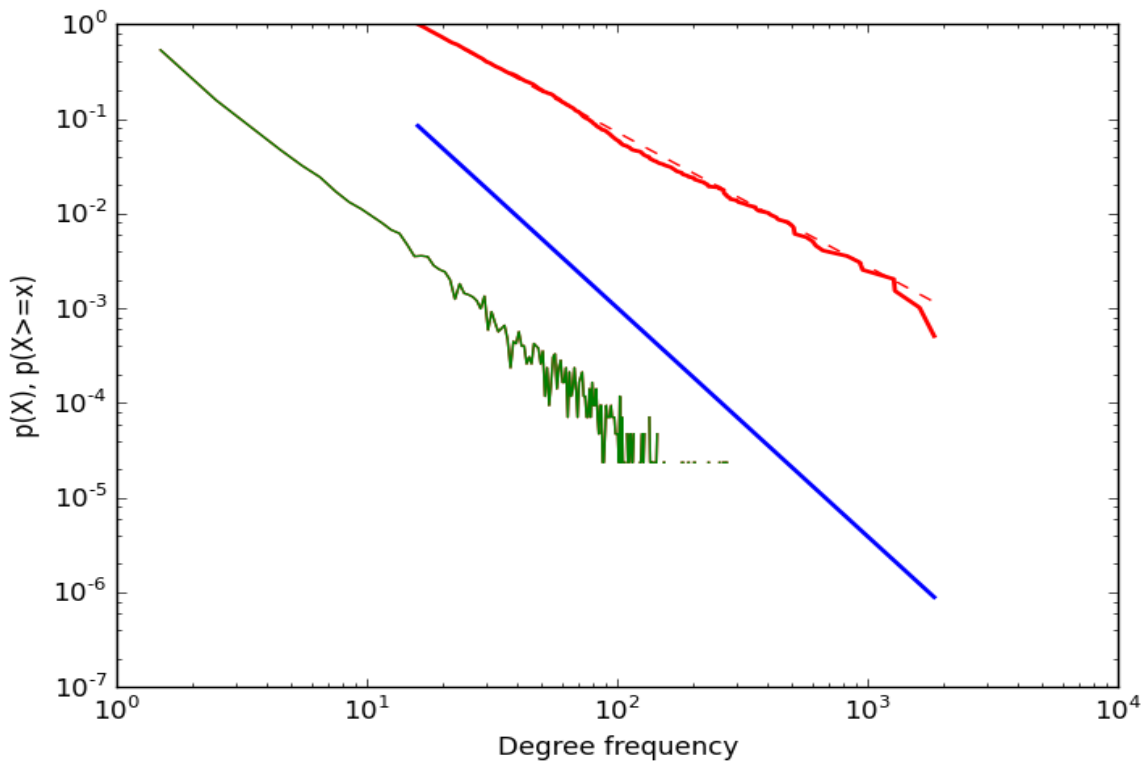
Results show that the probability density function (PDF) of a power law distribution fit ($\alpha=2.51, x_{\min}=42$) can well explain the degree distribution of the “contribution” network under study. The blue line in figure 52 with line style ‘-’ is the power law fit estimated on $x_{\min}=42$. The green line with linear-bins demonstrates the degree distribution on empirical data ($x_{\min}=1$). The noise data involves nodes with degree less than 42. This is consistent with arguments by Clauset *et al.* (2009) that the observed power law distribution with empirical data is often above the true x_{\min} . The upper red line plots the complementary cumulative distribution function (CCDF) that is known as the survive function or reliability function. CCDF captures the degree distribution follows the power law beyond the studied time. On log-log plot, CCDF can be written as: $\ln \Pr(X \geq x) = -\alpha(\ln x + c)$ (6.1) and PDF is: $\ln f(x) = (-\alpha - 1) \ln x + \ln \alpha - \alpha \ln C$ (6.2), where C is constant. Thus the plot PDF (-a-1) is on the left down

side of CCDF. CCDF can be expressed as: $1-F(x) = \frac{1-C \int_0^x (x^{-\gamma})' dx}{1-C} = \frac{(\frac{x^{-\gamma+1}}{-\gamma+1} - \frac{1}{-\gamma+1})}{1-C}$ (6.3), with

$(\frac{C}{\gamma-1})x^{-(\gamma-1)}$ indicating the slop of CCDF is slower than that of PDF. Because that $C = (a-1)x_{\min}^{a-1}$, the normalised power law distribution density function is (Newman, 2005):

$$p(x) = \frac{a-1}{x_{\min}^{-a+1}} \left(\frac{x}{x_{\min}}\right)^{-a+1} = \frac{a-1}{x_{\min}} \left(\frac{x}{x_{\min}}\right)^{-a} \quad (6.4).$$

Figure 37: Degree distribution



(Blue or middle: probability density function (PDF) estimates on optimal $x_{\min} = 18$; Green or lower: PDF estimated with empirical data; Red or upper: CCDF (complementary cumulative distribution function))

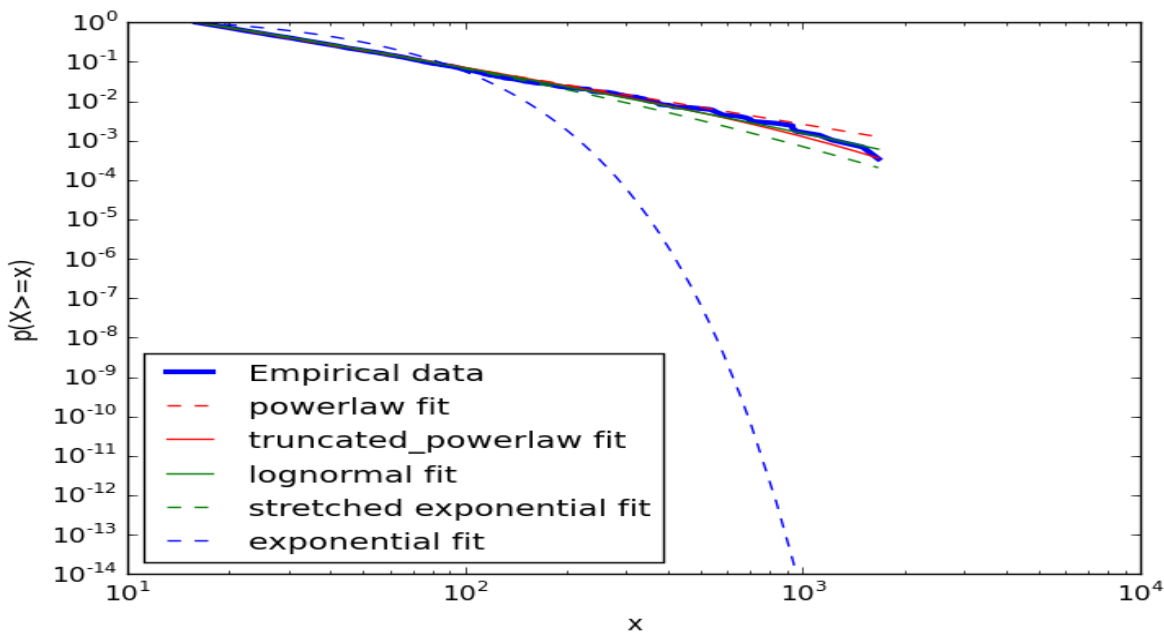
6.3.2 Comparing with candidate distributions

There are two generative mechanisms to scale-free networks, which can be summarized as continuing growth and preferential attachment (Barabasi and Albert, 1999). Continuous growth can encourage discussions within online forums that are demonstrated through the density of communications. Preferential attachment explains that new members are more likely to associate with popular experts

within online forums who have more connections than others. In other words, preferential attachment enables influential contributors to be more popular.

The mechanisms that generate a power law distribution are often compared with other distributions that are also in the heavy tail family. Exponential distribution that describes the random walk is the minimum comparative fit. Lognormal and stretched exponential distributions are other two possible candidates that can explain the preferential attachments (Mitzenmacher, 2003; Newman, 2005; Molontay, 2013; Alstott *et al.*, 2014). In this study, power law fit is in addition compared with truncated power law as there involves the lower boundary of minimum x_{\min} as discussed in the above power law fitting.

Figure 38: Comparing with candidate degree distributions



Stretched exponential distribution is also known as Weibull distribution, which is the form $e^{-(t/\tau)^\beta}$. When $\beta = 1$, it is the curve form for the exponential distribution. As $0 < \beta < 1$, the graph on log-log plot is stretched with $\log(x)$ but is a function of β which roughly follows normal distributions. Stretched exponential can be simplified as e^{-ax^b} , where a and b are constants. It is used to describe for instance the numbers of email

address books in a college that spans about three orders of magnitude but does not follow the power law distribution (Newman, 2005).

Lognormal distribution assumes random variables $Y_i = \ln X_i$, $i > 0$ have normal (Gaussian) distribution that has finite means and variances in contrast to power law distribution highlighting “richer getting richer” (Mitzenmacher, 2003). However, lognormal distribution is extremely similar to power law distribution on the log-log plot because its CCDF and PDF are straight lines, which makes it difficult to distinguish between them.

Embedded in the empirical data structure, simulations are taken on 10000 and 30000 randomly generated arrays in order to test the theoretical hypothesis, i.e. online contribution network is scale-free where a majority of knowledge contributions are held by a small part of members. Results show that truncated power law is superior to power law distribution in explaining empirical data (-0.0516, 0.7480). Power law fit explains better the empirical data than other competitive fittings.

Mossa *et al.* (2002) provide a possible argument in explaining the nested power law distributions observed with empirical data. In the BA model, preferential attachment explains that new members will join to existing members with a probability that is proportional to the numbers of edges to those members (Barabasi and Albert, 1999). This assumes that new members have information (which is called unfiltered information by Mossa *et al.*, 2002) about existing members. Although the Stack Overflow forum allows members being informed the numbers of followers associated with existing members, it is possible that new members do not notice such information for some reason (information processing capabilities available to members by Mossa *et al.* 2002). The information processing capabilities available to members can be the cause of the exponential truncation (Mossa *et al.*, 2002).

In general, the network effects contributed by a small group of contributors as a fraction of the total sampling (P) are a function of the scaling parameters (Newman, 2005): $Effects = P^{(\gamma-2)/(\gamma-1)}$. In this study, 1% percent of experts contribute around 25% of total effects; around 60% of ‘contribution’ network

connections are held by 20% of members; 80% of connections are fed by a half of members. If γ is little higher than 2, the situation is even more extreme. For instance, when $\gamma=2.1$, 20% of members will have around 86% of total connections, which is so called 80/20 rules (Newman, 2005).

Table 42: Power law fit generated by Python

	Simulated data(10000)	Simulated data(30000)	Empirical data (64597)	Support for power law
Xmin	42.003	42.7912	42.0	
alpha	2.5636	2.5819	2.5832	
Sigma	0.01564	0.00927	0.05587	
D	0.0071	0.00493	0.0142	
Power law vs exponential	(4562.7532, 0.00)	(15760.34387, 0.00)	(244.631, 3.141)	support
Power law vs stretched exponential	(72.344, 1.765)	(132.192, 1.6946)	(3.978, 0.0559)	support
Power law vs lognormal:	(-0.1690, 0.66316)	(0.2964, 0.0824)	(0.0153, 0.05526)	support
Power law vs truncated power law	(4.01992, 0.99284)	(7.60269, 0.991)	(-0.0516, 0.7480)	moderate support
Truncated power law vs lognormal	(-0.1691, 0.66192)	(0.293, 0.075)	(0.0669, 0.54524)	support
Truncated power law vs stretched exponential	(72.3434, 1.76696)	(132.193, 1.6973)	(4.0299, 0.04021)	moderate support

6.3.3 Self-sustaining scale-free network and Theory of critical mass

Previous studies in truncated power law can be understood as broken power law or power law with exponential cut off (Jóhannesson *et al.*, 2006) . Power law with exponential cut off (Clauset *et al.*, 2009)

is: $f(x) \propto x^\gamma e^{\beta x}$, with exponential curve down ward in the tails. With either situation, phase transitions are observed. In network analysis, truncated power law is expressed by the power law with exponential

cut-off (Newman *et al.*, 2003): $p(k) \propto k^{-\gamma} e^{-k_{\max}/k_{\min}}$ (6.5). For any finite scale-free network, there are lower and upper cut off (Cohen *et al.*, 2003): $p(k) \propto k^{-\gamma}$, where $k_{\min} < k < k_{\max}$. The highest degree $k_{\max} \cong k_{\min} N^{1/\gamma-1}$, thus k_{\max}/k_{\min} is the exponential expression of the network size (N) for $3 > \gamma > 2$ which is the regime for scale-free network (Barabasi and Albert, 1999). In other words, the power law with exponential cut off can describe the scale-free network in the real world (it can only be analysed with empirical data).

A phase transition is associated with the percolation theory that seeks to understand the cluster structure within a network (Cohen *et al.*, 2003; Newman, 2005). If nodes are randomly removed with a probability q , p is the probability of nodes remaining connected within a network. p_c is the critical point over which the network is connected with a large cluster, called giant component in a random network and a spanning cluster in a scale-free network (Cohen *et al.*, 2003). Spanning cluster is associated with an infinite system, which can describe the self-sustaining characteristics of scale-free network. A finite network which is connected for the first time is the critical point for the emergence of the spanning cluster.

The percolation phase transitions are geometric which involves the parameters such as network size and heterogeneity of systems. The probability that an arbitrarily chosen node belongs to the largest cluster size closing to the critical point is expressed (Cohen *et al.*, 2003): $p_{\infty} \propto (p_c - p)^{\beta}$ (6.6), where β is dependent of dimensions. The mean degree of the first neighbours is two times of dimensions ($\langle k \rangle = 2D$) (Barabasi and Albert, 1999), which is around 2.3 dimensions in this study ($\langle k \rangle \cong 4.6$). In scale-free

network, β can be calculated through $\frac{1}{3-\gamma}$ (6.7) ($1/(3-2.5832) \cong 2$ in this study). After p_c , the distribution

of cluster size is: $n(s) \propto s^{-\tau}$ (6.8), with $\tau = \frac{2\gamma-3}{\gamma-2} \cong 4$ in this study (Cohen *et al.*, 2003), suggesting a small

number of connected clusters or big size of cluster within the online forum.

Simulated data in the section of power law fitting shows that the network size plays a role in power law fit. The evolution of the network demonstrates phase transaction in term of network types. From the beginning, lognormal distribution is observed (power law vs lognormal), indicating that only a small part of members answer questions, and then there are isolated clusters because only a small part of members have connections. As more members join and participate in online discussions, the connections between members increase. An online forum is developed into a larger random network (power law vs exponential) with the increased probability for the emergence of giant components. Due to the preferential attachment, experts are more likely to answer questions, and they can have more and more connections over time. As a consequence of this, variances in terms of connections associated to members tend to be infinite over time in scale-free networks. Scale-free is the concept to describe the situation when the second moment is divergent (Barabasi, 2013).

However, the distribution of $p(s)$ is therefore dimensionless (e.g. Newman, 2005). The condition for the emergence of spinning cluster within indirect scale-free network is (Cohen *et al.*, 2003): $p_c = \frac{\langle k \rangle}{\langle k(k-1) \rangle} \cong 0.008$ which suggests a very stable network. As this criticality closes to zero, it can be considered without critical threshold. There is always a spinning cluster that tends to be infinite within online forums. It can also be understood that truncated power law with exponential cut off rather than the general power law is more suitable for describing the scale-free network analysed with empirical data (Newman *et al.*, 2001). That is, the online forum Stack Overflow is a random network only during the very early period, and it is self-sustaining because it has been developed into a scale-free network. Scale-free network is self-sustaining and it always percolates (spinning cluster always exists); in this sense, there is no threshold for scale-free network (Barabasi and Albert, 1999).

6.4 Evaluating the role of critical mass members

Critical mass members are contributors who have more connections during the network evolution. A degree centrality measure can identify the importance of members regarding their connections in the network. Other popular centrality measures, such as edge betweenness centrality, eigenvector and

PageRank are more suitable for understanding information diffusion within a network, because those centrality measures highlight the importance of neighbours to an arbitrary member.

The role of critical mass members in the contribution network is evaluated by examining the connectivity of an online forum after removing an increasing fraction of critical mass members. In other words, this is an attending attack to understand the critical point, denoting p_a in this study, after which online forum is no longer connected. Denoting p_r is the critical point to break down an online forum by random attack; a smaller p_a can reflect the importance of critical mass members.

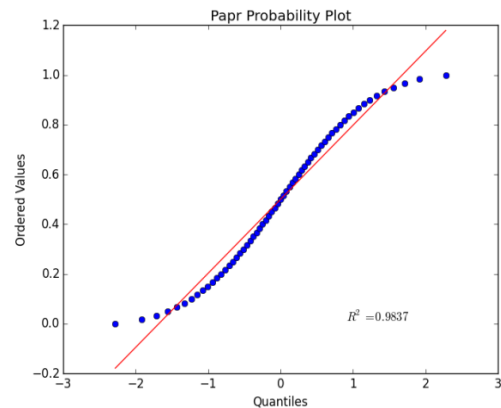
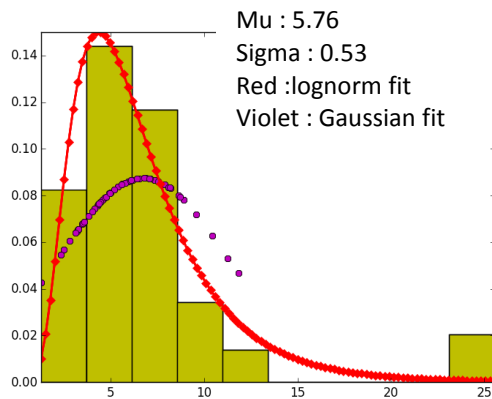
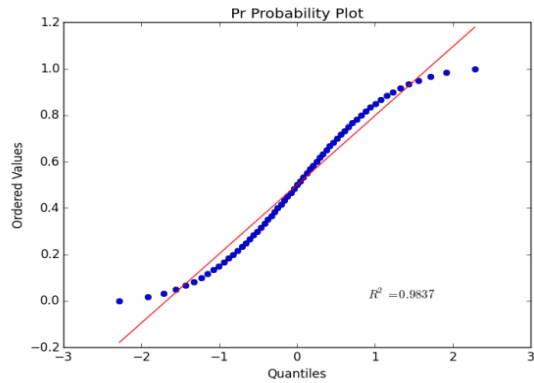
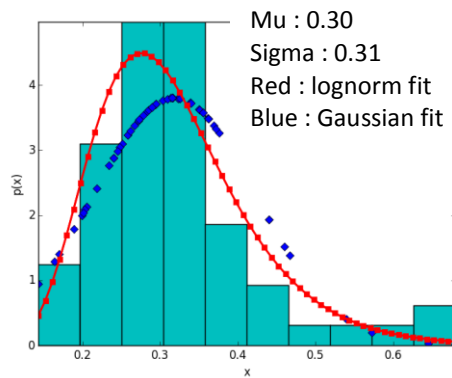
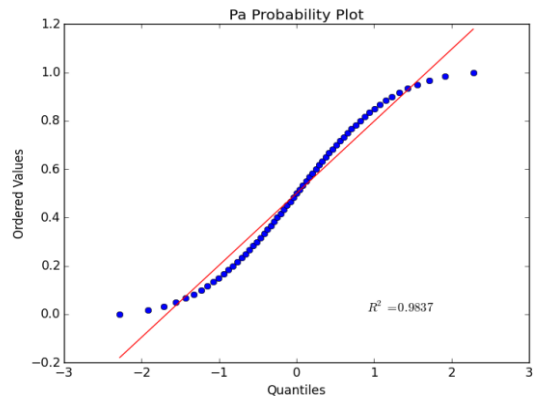
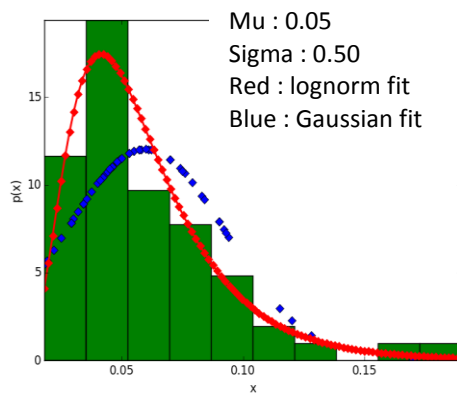
To evaluate p_a and p_r , simulation is firstly performed on NetworkX for Python using the empirical data characteristics, i.e. giving alpha (=2.58), average clustering coefficient (=0.023) and mean degree 5 (4.6 round to 5 for an integer). Restrained by the author’s computer memory, simulations are only repeated 60 times on different network size ranged from 100 to 2000. The following table describes the data collected using simulations.

Table 43: Attending and random attack

	Attending attack p_a			Random attack p_r			PrPa		
size	group1	group 2	group3	group1	group 2	group3	group1	group 2	group3
100	0.019	0.018	0.029	0.44	0.46	0.68	23.15789	25.55556	23.44828
200	0.19	0.17	0.12	0.24	0.54	0.645	1.263158	3.176471	5.375
300	0.083	0.07	0.093	0.467	0.573	0.44	5.626506	8.185714	4.731183
400	0.043	0.115	0.128	0.315	0.323	0.295	7.325581	2.808696	2.304688
500	0.092	0.09	0.084	0.35	0.314	0.372	3.804348	3.488889	4.428571
600	0.055	0.058	0.062	0.325	0.285	0.203	5.909091	4.913793	3.274194
700	0.061	0.053	0.046	0.351	0.376	0.201	5.754098	7.09434	4.369565
800	0.076	0.074	0.075	0.189	0.316	0.309	2.486842	4.27027	4.12
900	0.049	0.048	0.07	0.17	0.24	0.362	3.469388	5	5.171429
1000	0.094	0.05	0.079	0.303	0.326	0.315	3.223404	6.52	3.987342
1100	0.048	0.09	0.06	0.361	0.295	0.362	7.520833	3.277778	6.033333
1200	0.034	0.055	0.029	0.294	0.251	0.278	8.647059	4.563636	9.586207
1300	0.046	0.042	0.04	0.339	0.258	0.272	7.369565	6.142857	6.8
1400	0.03	0.029	0.025	0.355	0.258	0.218	11.83333	8.896552	8.72
1500	0.041	0.032	0.046	0.323	0.205	0.143	7.878049	6.40625	3.108696

	Attending attack p_a			Random attack p_r			PrPa		
1600	0.044	0.046	0.046	0.323	0.199	0.164	7.340909	4.326087	3.565217
1700	0.057	0.042	0.044	0.369	0.34	0.245	6.473684	8.095238	5.568182
1800	0.034	0.057	0.053	0.355	0.267	0.262	10.44118	4.684211	4.943396
1900	0.034	0.022	0.045	0.234	0.248	0.273	6.882353	11.27273	6.066667
2000	0.0375	0.0355	0.041	0.283	0.289	0.289	7.546667	8.140845	7.04878

Figure 39: Probability plot of network attacks



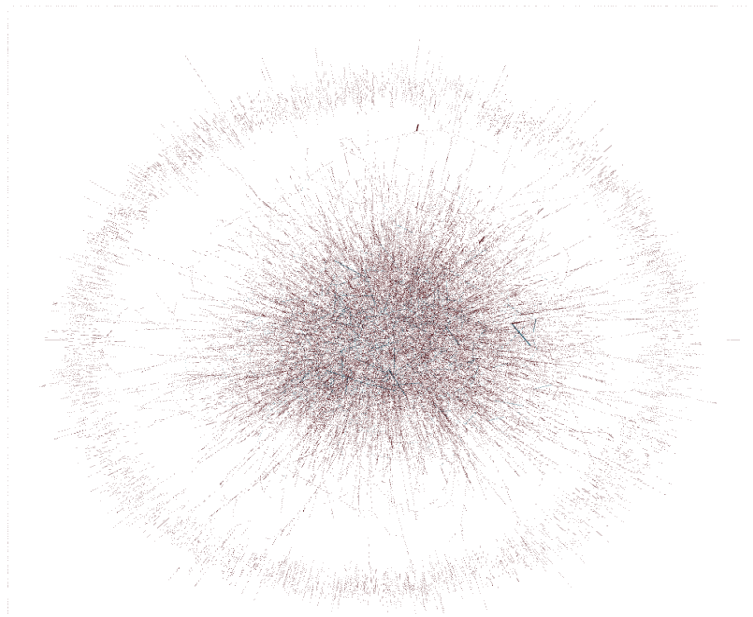
The distributions of p_a and p_r roughly follow the lognormal distributions. The probability CCDF distribution fit for p_a , p_r and p_a/p_r suggest $\ln(x)$ can roughly fit normal distribution with squared R closing to 1. In general, to remove of around 5% of hubs can break down online forum stack over flow. However, it requires to randomly removing 30% (around 6 times) of nodes so that the scale-free network is disconnected.

Simulated results are consistent with those generated from the study by Albert *et al.* (1999) who firstly explored the robustness (tolerance to random error) and fragility (intolerance to attending attack) characteristics of the scale-free network. By removing 7360 experts (147190×0.05) who contribute around 33.33% of knowledge ($P^{\frac{\gamma-2}{\gamma-1}}$), an online forum is broken down. The following figure 54 shows disconnected small clusters located in perimeter after attending attack.

Before an attending attack, the average distance scales between any of two members within an online forum at $6.03 \left(\frac{2 \ln \ln(N)}{|\ln(\gamma-2)|} \right)$ (Cohen and Havlin, 2003)). The average distance between members increases to 14.722 after the attending attack, i.e. removing an increasing proportion of critical mass members. Wang and Dai (2009) define that the network efficiency equals to 1 divided by the average distance between members: $\frac{1}{\bar{d}_{ij}}$. That is, the network efficiency decrease around 59% by removing around 5% of critical mass members $\left(1 - \frac{(1/14.722)}{(1/6.03)} \right)$.

In summary, critical mass members play an important role in the evolution of online forums. Not only do they pay the digital public goods contribution cost, but also determine the expansion of the network in size as well as the communications within the online forum.

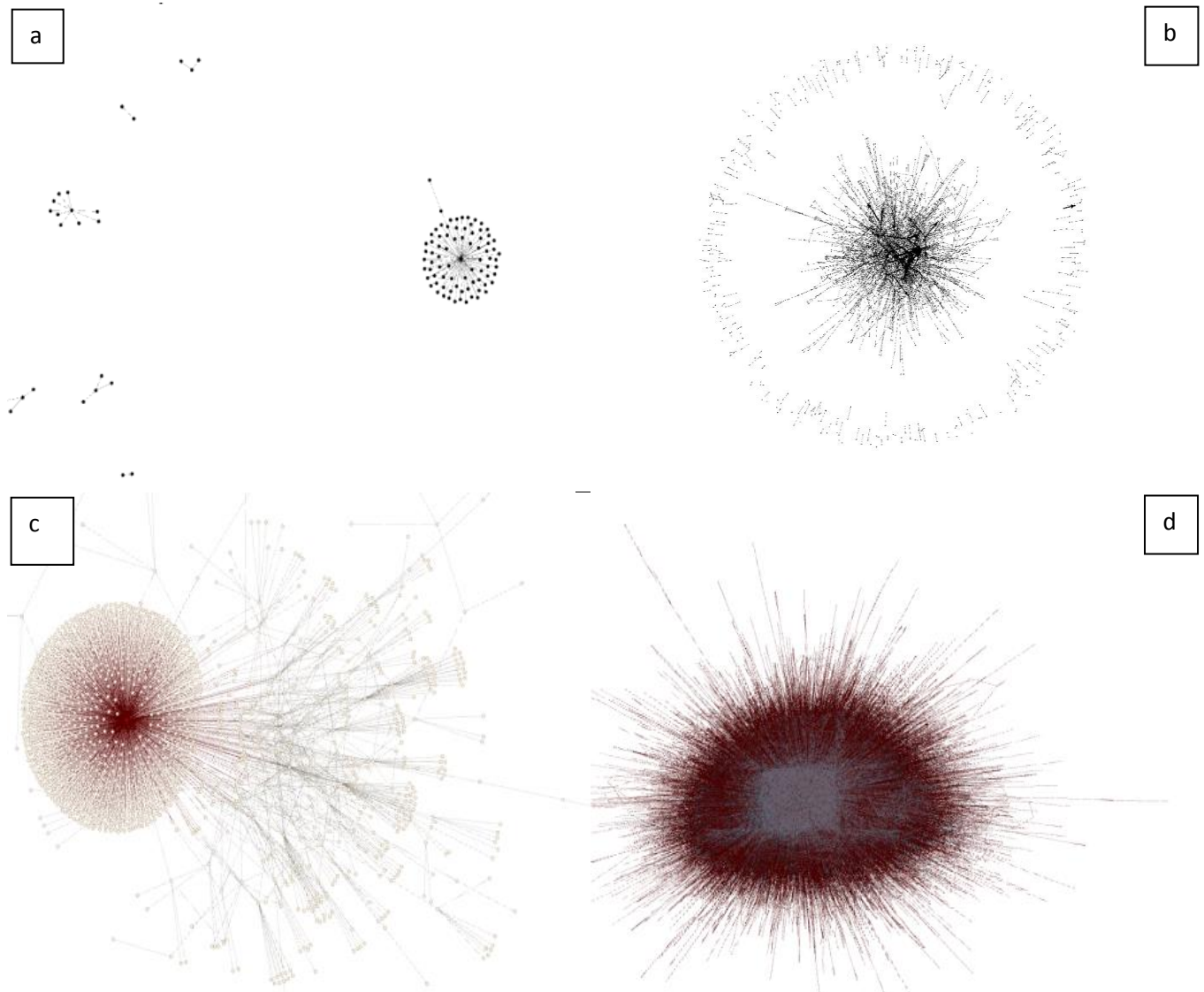
Figure 40: After removing 5% of critical mass members



6.5 Evolution of online forums

Using the data visualisation software Gephi, the following describes the evolution of an online forum embedded in the empirical data. Under criticality, there are isolated small clusters. Closing to the criticality, the largest cluster size is much larger than that in the very beginning. Above the criticality, it is showing that the largest cluster tends to grow in the future (spinning trees are coloured in red). Spinning trees are edges (bonds) that connect nodes in the largest cluster to others who do not yet belong to the same cluster. It also shows that members tend to connect with experts who are more likely to answer questions and therefore who have more connections, which results in those experts having more and more connections over time. This is consistent with the argument by Barabasi and Albert (1999) that the majority connection in a scale-free network is held by hubs. In addition, experts are more likely to connect with each other (so called assortative matching/mix), because they are more likely to exchange ideas around a topic in online discussion forums.

Figure 41: The evolution of online forum

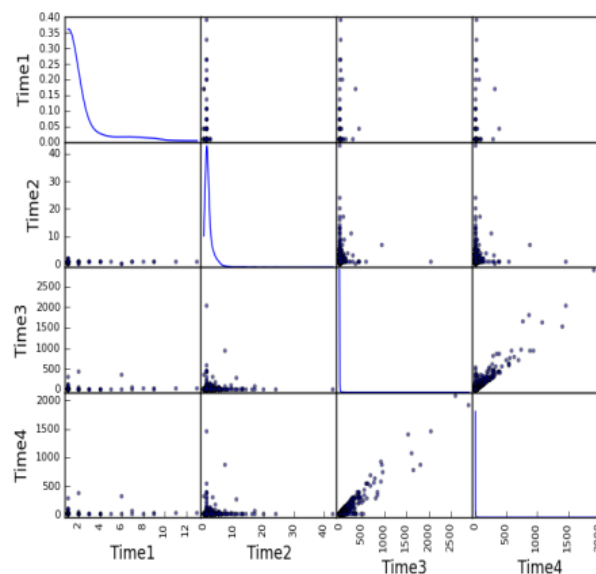


- (a) Capture of the situation far before the criticality (time1), where the online forum is disconnected; (b) Capture of the moment closing to the critical point (time2), where the largest cluster size is much bigger than before. At this moment, the online forum is still disconnected. (c) and (d) capture the situation after the criticality around (time3 &4), where the online forum is connected with a spinning cluster that is only limited by the network size. (c) Shows that a small group of nodes have more connections than others, and the online forum has demonstrated the important property of scale-free networks. (d) Captures the network far above the criticality, spinning trees in red suggest the continuous growth.

The above discussions are consistent with results generated from study one where the concept of perceived critical mass within online forums are measured through the perceived density of communications and perceived linkages with some others. Digital public goods are often initially contributed by small fractions of members who pay the start-up cost. Above a critical point, the size of online forums tends to be infinite because new members attracted by knowledge provided by initial

contributors continuously join online discussions. If discussions are not interesting, members may leave. In social science, the interests of a topic are demonstrated through the audience (Westland, 2010). Therefore, initial members are important for the development of online forums, not only because they have a willingness to contribute but also they are more knowledgeable. This is in agreement with the production function of critical mass model, which explains that the contribution behaviour is more likely to happen when resources allocated between members are heterogeneous. In return, initial contributors win more and more connections (social capital), which suggest the network structural influence on their motivation to contribute. The theory of critical mass (Olivier and Marwell, 1988) that involves a mass contribution evoked by a small group of contributors is supported in the context of online forums.

Figure 42: Probability density functions of the 4 time periods



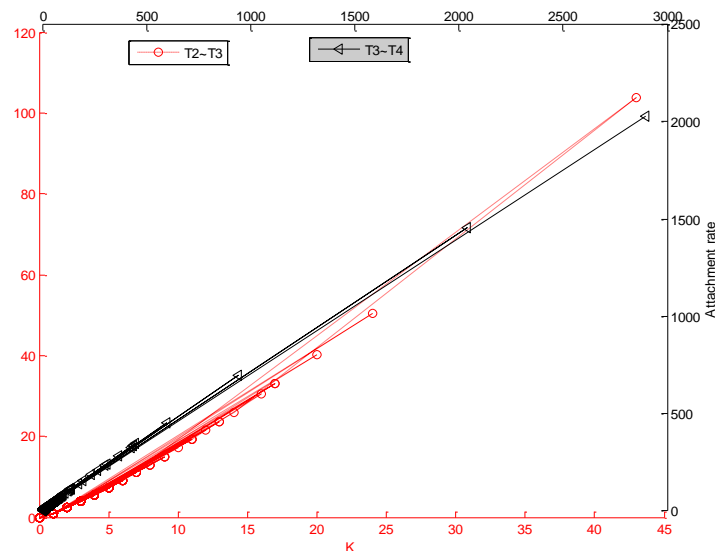
The diagonal of the scatter matrix plots the probability density function (PDF) of each stage, with its own conjugate transposition comparing with other stages.

There are different network growth models in terms of the preferential attachment between members, measured with the parameter (a) (e.g. Jeong *et al.*, 2001; Kunegis *et al.*, 2013). When “ a ” equals to 1, it suggests linearly attachment and corresponds to the BA model. When “ a ” exceeds 1, it refers to the super linear preferential attachment within which a small fraction of ‘winner takes all’, i.e. the numbers of hubs

are limited particular for the super-linear case. When “ α ” is smaller than 1, it is called the sub-linear attachment, and the degree distribution of the network is often a stretched exponential (not suitable for study three) (e.g. Jeong *et al.*, 2001).

The super-linear preferential attachment is found as a remarkable mechanism that governs the expansion of a network. This again highlights the role of the initial hubs because the newly joined members will create links to the existing members based on their connectivity. However, after the critical point, the numbers of hubs increase because members tend to linearly attach to hubs. This is in agreement with the previous findings (e.g. Barabasi, 2013) that linear attachment is the mechanism in the region of scale-free network. In other words, seeding hubs in a later stage will increase the cost of seeding.

Figure 43: Preferential attachment plot of the successive stages



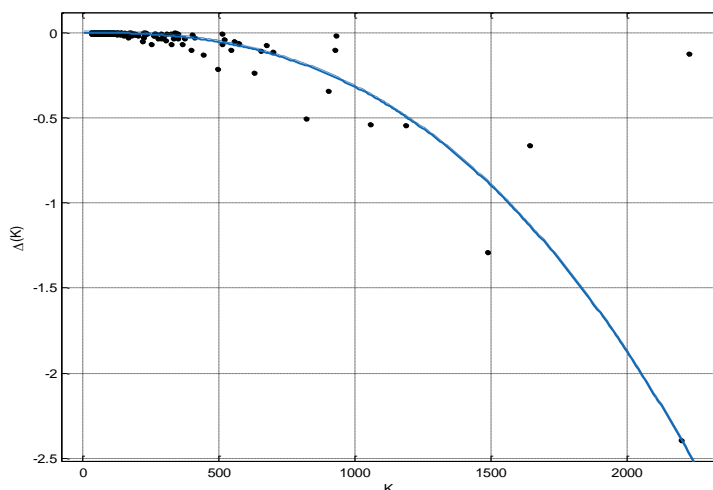
(T2~T3: before to after the critical point with $\alpha=1.2342$; T3~T4: after the critical point with $\alpha=0.95537$)

Moreover, it is noted that the highest degree decreases in the time period 4 when comparing with that in the time 3. This is in agreement with previous studies (e.g. Clauset *et al.*, 2009) that high degree nodes have a tendency to drop down in their connections, which leads to the exponential cut-off in the tail of the degree distribution. One possible explanation of this result can be, with the expansion of the population,

people are able to gather in different groups rather than in a single group (Small *et al.*, 2014), and hence it limits the size of hubs.

Although hubs identified in each stage are not necessarily duplicated, it is observed that one member with the highest degree presents in the four successive periods. Barabasi (2013) argues that the probability of connecting to the old nodes is higher than to the young nodes.

Figure 44: Dispersion of degree plot



The fluctuation is observed for hubs, while the dispersion of the hub (top right dot) with the highest degree is relatively stable.

In summary, structural influence on members' intention to contribute within online forums can be explained using the properties of scale-free network. The scale-free network is self-sustaining (Newman, 2003; Barabasi, 2013) and it allows contributors to be more popular. For an online forum to be successful, it is important that it is developed into a scale-free network after a critical point. That is, it should have some initial knowledgeable contributors who can attract more members to participate in online discussions. This study so far has provided rich information on the structural influence on members' contribution behaviours, and linked findings with the theory of critical mass that also highlights the importance of initial contributors in phase transactions. Studies of the structural influence on online

contribution behaviours are rare (Reedings and Wasco, 2012), and this highlights the theoretical contribution of this study. With managerial considerations, maintaining influential members is more important than satisfying everybody, because 80% of network effects are contributed by less than a half of members, and only 1% of hubs who have many connections can contribute 25% of network effects as an example raised with Stack Overflow online forum.

6.6 Conclusion

The structural influence on online contribution behaviours is evident. An online forum is characteristic of a scale-free network, where a small group of members hold majority connections and ensure the evolution of the network (Barabasi and Frangos, 2014). A phase transition is observed in dynamic aspects, with the critical point over which a brutal exposition with a different exponential growth rate. The scale-free structural influences also means the tolerance of free riders, a digital public good can be created by critical mass members. In return, critical mass members can gain social capital (connections and reputation) and this is one possible explanation of structural influence on contribution behaviours.

Increasing attention has been given to the identification of seeding targets within social media, with several notable studies (e.g. Watts and Dodds 2007; Hinz *et al.*, 2011; Libai *et al.*, 2013). It has been observed that published empirical studies using real network data remain limited (Hinz *et al.*, 2011; Libai *et al.*, 2013). This study addresses this point by using real life data obtained from a large-scale, successful online forum.

Although previous studies have shown that the selection of seeding candidates within social media is best studied using network analyses techniques (e.g. Hinz *et al.*, 2011), limited empirical studies have sought to use a network topology approach. Indeed, results from computer simulations have suggested that the implementation of a seeding program is influenced by the structure of networks (e.g. Bampo *et al.*, 2008). Schulze *et al.* (2014) who use regression analyses techniques in their empirical study find that the network effect is one contributing factor to the success of information diffusion. Despite the importance of the network structural influences on the seeding programmes, few empirical studies have sought to

take a dynamic view by investigating how the network develops (e.g. Shmueli and Altshuler 2014), and is associated with the theory of critical mass (e.g. Centola, 2013). This study contributes to the literature by providing additional insights to network evolution in the consecutive stages within which seeding programmes can be optimised.

Conventional wisdom in marketing is based on the “influential” hypothesis that “opinion leaders” have disproportional effects on their followers in the formation of public opinion (Katz and Lazarsfeld, 1955; Lazarsfeld *et al.*, 1968). Much attention has been given to the selection of the opinion influencers (Libai *et al.*, 2013). One research stream suggests that “opinion leaders” within a network, referred to also as “influencers”, “well-connected-members”, “hubs” or “high-degree-seeding”, are the seeding objectives who can ensure the rapid spread of information, because seeding ‘hubs’ can lead to a higher level of referrals (Bampo *et al.*, 2008; Hinz *et al.*, 2011). However, another research stream (e.g. Watts and Dodds, 2007) suggests that seeding hubs is not systematically a more efficient method; the success of a seeding program is more likely to rely on the characteristics of hubs such as being ‘easy-to-influence’. Because members are ‘easy-to-influence’, word of mouth is more likely to be an influential factor in the initial stage so that the critical point highlighted in the study by Watts and Dodds (2007) can be achieved quickly. The study by Libai *et al.* (2013) indicates a similar conclusion, i.e. word-of-mouth accelerates the expansion of networks. Referrers are weakly-connected-members in a network (Libai *et al.*, 2013), and the effects generated from “referrals” have been found to be efficient for spreading word-of-mouth about a brand (Kaikati and Kaikati, 2004). This highlights the social value of an individual customer in a seeding program and several terms have been used to measure such social effects, including “indirect effects”, “referral value”, “influence value” and “social value” (Libai *et al.*, 2013). In summary, previous studies led to a variety of conclusions about optimal seeding targets. It is worth noting that the above debate is fundamentally about whether the network is able to be developed into a stable state after a critical point and what role the hubs play within it.

The findings from this study can be one possible explanation of the diversity of findings from prior studies. For the final stage of the overall development, results suggest that hubs do play an important role in information diffusion once a network is scale-free. More precisely, it is found that maintaining only around 5% of hubs can ensure the information through the whole network; even if it is large scale in size. This result is in agreement with the distinction made by Albert *et al.* (2000) that a scale-free network is both robust to random attacks and fragile to attending attacks.

However, the network topology suggests that hubs always exist within a scale-free network (e.g. Newman, 2005). As a result of this argument, there will be no need to maintain hubs as long as it is a scale-free network. The examination of the evolution of the network shows that the majority of hubs are not necessarily duplicated over time. Although super-linear preferential attachment plays a remarkable role in network expansion (i.e. super-hubs through whom the seeding messages can pass through the entire network present before the critical point), new hubs are observed emerging in the later stage, and only the one who has the highest connections is present from the beginning. Moreover, the increasing rate of the numbers of connections to hubs tends to slow down over time. These two findings suggest that the probability of the participation rate of the “old” hubs is higher, in agreement with the argument by Barabasi (2013), while, the conclusion that the initial emerged hubs are more active in the online activities (e.g. Hinz *et al.*, 2011) should be cautious.

In the context of online forums, members are popular because they are more likely to have knowledge to share. Thus, it is argued that factors such as word of mouth and content quality may play a more important role than seeding hubs in the very beginning of a seeding program because those non-monetary factors can explain why the critical point after which the networks are self-sustaining is achieved quickly.

There are several managerial implications of the findings which could apply specifically to firm-hosted online forums or more generally to online social media networks. Marketers should design their network seeding programmes according to the stage in the evolution of a network. The objective is to have a firm-

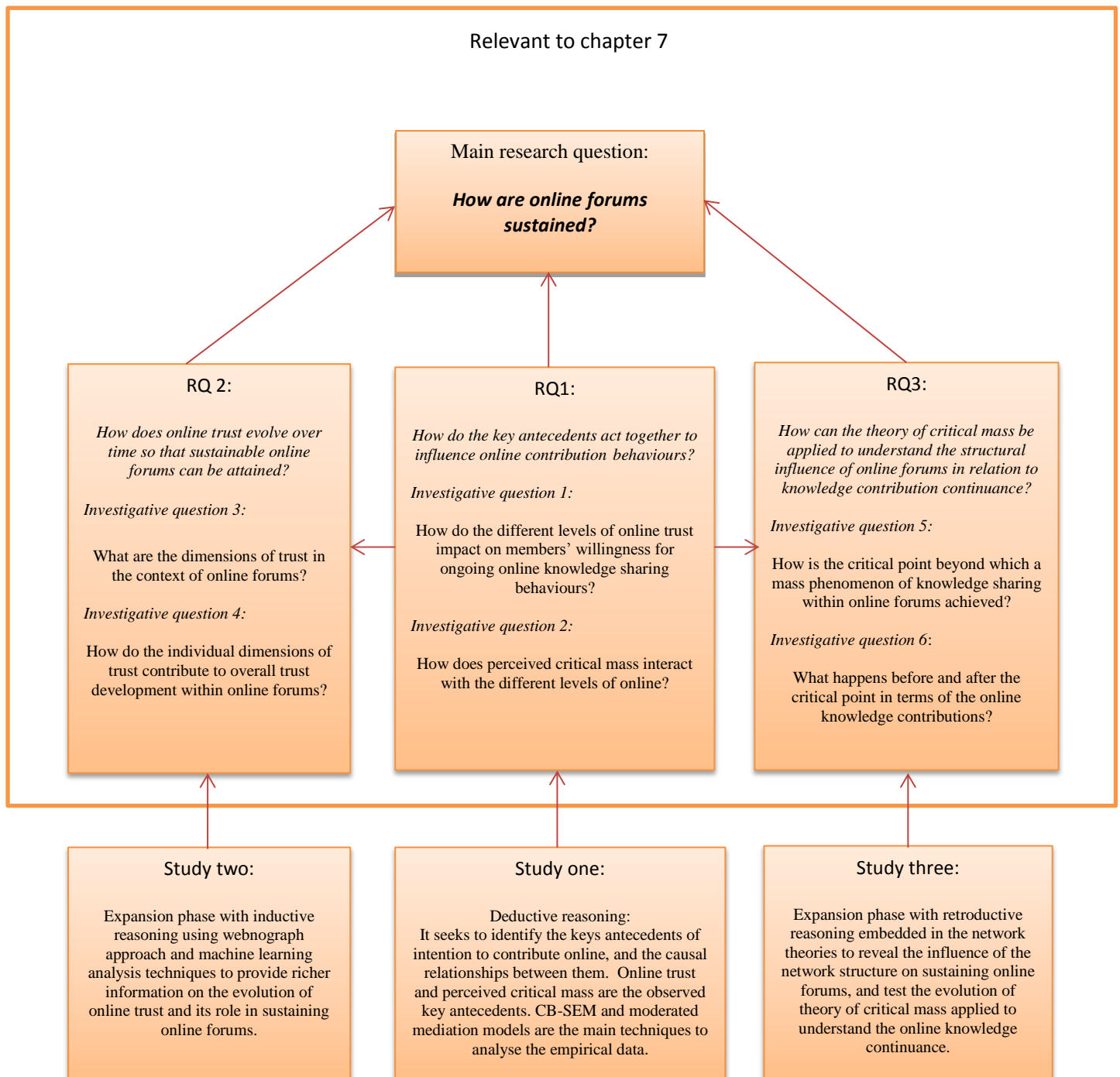
hosted social media network which is scale-free. Because the initial stage of the network development involves random factors and is relatively stable once it is scale-free, how to attract the earlier adopters of online forums seems to be the most important concern in the beginning. A majority of previous studies proposes the use of incentives (e.g. Schulze *et al.*, 2014), but this can be expensive in particular for small-medium size companies. An efficient seeding method can lead to the reputation of a network building through word-of-mouth during the implementation stages. In this stage, companies should try to have experts participating in online forums, because those experts would be more likely to provide the attractive content that could promote the reputation building and spread word-of-mouth. When the network is developed into a stable stage, maintaining hubs should be the essential task for marketers, because missing hubs (only 5% of memberships) would lead to failure in information diffusion. This suggests that companies can develop their capability in terms of hosting numbers of members so that hubs can be more and more popular. This may be achieved, for example, by having voting systems that encourage members to publish quality content, having limited moderation of the exchanges between established members and seeking to have thousands or millions of members online simultaneously. In summary, marketers are encouraged to consider the network effects as an important factor in their portfolio management.

This study is limited in different ways. Firstly, Stack Overflow is an example of a successful online forum, while what has happened in dying online forums is little discussed. Secondly, although examining the connections to members is the method adopted by a majority of previous studies, alternative methods may lead to different results in the identification of “influencers”. For instance, PageRank and HIT are diffusion algorithms that have a focus on the importance of neighbours and the importance of hubs in the process of information diffusion may be either overestimated or underestimated. Thirdly, study three have a focus on the influence of network effects, and do not test other contribution factors.

Chapter 7: Conclusions

Figure 7.1 reminds the rationale of this thesis:

Figure 7.1: Revisiting the rationale of this thesis



7.1 Introduction

This thesis set out to explore the sustainability of online forums. A sustainable online forum is characterized by the constant knowledge on a particular topic contributed by members of that forum over time (e.g. Harris and Rae, 2009). Online forums are increasingly important for businesses, representing an additional interface with their customers and suppliers (Füller *et al.*, 2008; Dholakia *et al.*, 2009; Demange, 2010). Typically, online forums can be a source for new product ideas, or may be useful in resolving queries that otherwise need to be resolved by the company's (paid) employees (Demange, 2010). However, an online forum would not be sustainable without the availability of knowledge from members (Chiu *et al.*, 2006; Harris and Rae, 2009; Wasco *et al.*, 2009).

Yet, research that has sought to investigate the dynamic antecedents of online knowledge contribution behaviours is rare (e.g. Chen, 2007). Existing studies tend to focus on isolated factors that have an impact on online knowledge sharing behaviours (e.g. Chiu *et al.*, 2006; He and Wei, 2009; Shen *et al.*, 2013). Moreover, questions such as how these identified antecedents act together to influence online intentional contribution behaviours, and how these antecedents developed overtime and play a role in sustaining online forums, remain open to be further investigated.

This thesis has firstly identified the important dynamic factors (trust in members, trust in online forums and perceived critical mass) that have an impact on the intention to contribute knowledge online. It proposed an integrative model and captures the causal relationships among the key influential factors, demonstrated with study one. Embedded in the previous studies, online trust (e.g. Wasco and Faraj, 2005; Chen, 2007; Zimmer *et al.*, 2010) and perceived critical mass (e.g. Shen *et al.*, 2013) were identified as the key antecedents of ongoing online contribution behaviours. To further investigate the dynamic aspects of these antecedents, study two was designed to examine the evolution of online trust and study three has sought to investigate the phenomenon of critical mass developing over time.

Previous theoretical literature that crosses diverse disciplines in answering the vital research questions to the topic of sustainability of online forums is inconclusive and rarely integrated. The three empirical

studies were designed to understand this topic by taking a holistic approach to answer the following three research questions:

RQ1 How do the key antecedents act together to influence online contribution behaviours?

- Investigative question one: how do the different levels of online trust impact on members' willingness for ongoing online knowledge sharing behaviours?
- Investigative question two: how does perceived critical mass interact with the different levels of online trust?

RQ2 How does online trust evolve over time so that sustainable online forums can be attained?

- Investigative question three: what are the dimensions of trust in the context of online forums?
- Investigative question four: how do the individual dimensions of trust contribute to overall trust development within online forums?

RQ3 How can the theory of critical mass be applied to understand the structural influence of online forums in relation to knowledge contribution continuance?

- Investigative question five: how is the critical point beyond which the mass phenomenon of knowledge sharing within online forums achieved?
- Investigative question six: what happens before and after the critical point in term of the online knowledge contributions?

Responses to the investigative research question one were able to identify the key antecedents of ongoing online knowledge sharing which is the intrinsic reason for sustaining web-based discussion groups, and the causal relationships among them. Both research questions two and three were designed to expand upon the antecedents identified and embedded in research question one accordingly. The study design

allowed the research to be cross disciplinary and integrative, thereby addressing a lack of understanding in this field.

7.2 Empirical findings

To address the above three research questions, which relate to the main research question of how to sustain web-based discussion groups, three empirical studies were undertaken. Study one took the deductive approach using an online survey to collect data. Results of study one were further expanded in studies two and three. Study two used Webnography and machine learning techniques to understand the evolution of the overall online trust and its dimensionality. Study three followed the network structural analysis and has described how the critical mass members play a role in sustaining online forums. Studies two and three took a dynamic view, and complemented study one that was cross-sectional. Results generated from the three empirical studies were chapter specific and summarized in chapter four, five and six accordingly. This section will synthesize the findings from the empirical studies in answering the research questions.

RQ1: *How do the key antecedents act together to influence online contribution behaviours?*

- Investigative question one: How do the different levels of online trust impact on members' willingness for ongoing online knowledge sharing behaviours?
 - Investigative question two: How does perceived critical mass interact with the different levels of online trust?
- i. Knowledge sharing is an intentional behaviour: intention can reflect behaviour (Ajzen, 1991). Decomposed theory of planned behaviour (Tylor and Todd, 1995) is suitable for understanding the motivational factors of voluntary contribution online.
 - ii. An integrative view should be taken to understand online knowledge sharing behaviours: previous studies have identified isolated or limited numbers of online voluntary contribution behaviours embedded in different theoretical approaches. Hypotheses developed in study one

which incorporates dynamic antecedents were tested using two-step CB-SEM. Overall findings showed that intention to online knowledge sharing was directly determined by attitude, perceived control behaviour and subjective norms. The antecedents of determinants identified in this study are in agreement with previous studies that interpersonal trust (Wasco and Faraj, 2005), institutional trust (Chen 2007; Erden *et al.*, 2012) and perceived critical mass (Shen *et al.*, 2013) are the key dynamic predictors of online knowledge contribution continuance. However, the antecedents of drivers and their different magnitude influential levels have not been examined simultaneously in existing studies. It is the study's rationale in which there is a need for an integrative view when understanding online knowledge sharing behaviours.

- iii. Trust in online forums has more weighted power in predicting online intentional contribution behaviours than trust in members and perceived critical mass do: perceived linkage to the critical mass members and perceived size of contributors within online forums were essential in normative intention (i.e. H8 that supposed the perceived critical mass affecting on the subjective norms was supported), in agreement with previous findings that the perception of the mass contributions evoked by critical mass members can give members social pressure with regard to their intention to contribute online (e.g. Cho, 2011). Trust in members positively impacted on “subjective norms” which was one antecedent of intention to contribute online (i.e. H7 that hypothesised trust in members impacting on the subjective norms was supported). This is in agreement with the arguments by Jeffries and Becker (2008) that high levels of trust in others may lead to social influences on intended behaviours. Trust in online forums played a role as the contextual factors influencing “perceived behavioural control” and “attitude”, which were the others two determinants of online intentional behaviours (i.e. H4b hypothesised trust in online forums positively affecting on the attitude, and H6 supposed trust in online forums positively influencing on the behavioural control were supported). Zimmer *et al.* (2010) find that trust in online communities can lead to the positive attitude to purchase online; Erden *et al.* (2012) find that trust in the competences of online communities is positively associated with perceived

behavioural control. Given the fact that trust in online forums has a positive effect on the two determinants of online intentional behaviours, it was concluded that trust in online forums has greater magnitude levels of influences on online voluntary contributions, if it is compared with trust in members and perceived critical mass.

iv. Trust in members who are not behaving opportunistically favour the development of trust in online forums: trust is multidimensional and can be studied with different levels (e.g. Ridings *et al.*, 2002). Embedded in the proposition by McKnight *et al.* (1998), trust in members refers to interpersonal trust, while trust in online forums is associated with the more general institutional trust. Findings from study one showed that trust in members could lead to trust in online forums (i.e. H5 that supposed trust in members impacting on trust in online forums was supported), in agreement with the argument that interpersonal trust can lead to institutional trust (e.g. Luo, 2006). Moreover, study one further clarified that the affected based (i.e. benevolence and integrity) interpersonal trust were the antecedent of institutional trust, which is consistent with the argument by Levin *et al.* (2003) that affective based interpersonal trust in the context of knowledge sharing is associated with the perceived knowledge available within an organisation, and the knowledge available can reveal the ability of that organisation.

v. Trust in online forums completely mediates the effects of trust in members to attitude: the path from trust in members to attitude was not found to be statistically significant (i.e. H4a that supposed trust in members influencing on attitude was not supported), which disagrees with the arguments that social recognition is a predictor of attitudinal knowledge sharing behaviours (e.g. Jiang *et al.*, 2002), and that social recognition is both the antecedent and consequence of interpersonal trust (e.g. McKnight and Chevery, 2002). However, examination of the causal relationships between trust in online forums, trust in members and attitude showed that trust in online forums completely mediated and moderated the attitude regressed on trust in members. This result was in agreement with the support of H5 (i.e. trust in members could lead to trust in online forums). Thus, the effects of trust in members could not be ignored but the role of trust in online forums was highlighted.

- vi. Perceived critical mass partially mediates the effects of trust in members regressing to subjective norms: embedded in the propositions by Granovetter (1973) and Haythornthwaite (2002), trust in members at the individual levels that reveals the reciprocal frequency and homophile among members was understood to be associated with strong ties; perceived critical mass and trust in online forums that occur within a wider communications were considered linking to weak ties. Trust in members was found to positively affect subjective norms (i.e. H7 was supported); it was further explained that the effect of trust in members on subjective norms were partially mediated by perceived critical mass. However, such mediation effects disappeared when they were tested within the integrative model proposed by the study. In other words, the findings showed that weak ties could influence strong ties in the context of text-based online knowledge sharing. Tests of the measurement models showed that critical mass members often play the role of “bridge” between communications among members, in agreement with the argument by Granovetter (1973) that weak ties can be the linkages between groups involved in strong ties, and can enlarge the communications among individuals.
- vii. Trust in online forums completely mediates the effects of perceived critical mass regressed on trust in members: trust in members was not found to significantly affect perceived critical mass (i.e. H9a was not supported). However, a causal examination showed that the relationships between trust in members and perceived critical mass were completely mediated by trust in online forums. This result has highlighted the strength of weak ties; again it was in agreement with the idea that weak ties can span the numbers of interactions and promote an expansion of membership (e.g. Granovetter, 1973; Centola, 2013).
- viii. Strong and weak ties co-exist and play a role in online knowledge continuance: although the above discussions in (v), (vi) and (vii) have highlighted the importance of weak ties within online knowledge contributions, the role of strong ties cannot be ignored, which was demonstrated by the support for H5 (see discussions in (iv)). That is, the previous findings generated with simulations (i.e. Centola, 2013) that strong ties can impede the expansion of memberships were questioned with the findings from study one.

RQ 2: How does online trust evolve over time so that sustainable online forums can be attained?

- Investigative question three: what are the dimensions of trust in the context of online forums?
 - Investigative question four: how do the individual dimensions of trust contribute to overall trust development within online forums?
- i. Online trust evolves over time: the majority of previous research has not investigated the dynamic aspect of online trust (e.g. Palmer and Huo, 2013). Results generated from the neural network for time series analysis (NNT) indicated that, although trust in a popular brand tends to increase overtime, the overall scores of online trust vary in different stages. Data were collected from 2005 to 2010. Results showed that the scores for overall trust indicated a slow rising pattern for the year 2006, followed by a more rapidly falling pattern in 2007 to 2008. The overall trust score rose in the year 2009 and decreased again in 2010.
 - ii. The concept of online trust has three dimensions, which are ability, benevolence, and integrity: embedded in comments about a brand left by reviewers from three different online forums, study two found that the concept of online trust has four dimensions, which were ability, benevolence, integrity and predictability. Results from support vector machine (SVM) showed that all of those components were distinctive from each other. However, predictability contributes least to the concept of online trust, in agreement with findings by Lu *et al.* (2009).
 - iii. Individual dimension of online trust plays different roles in the evolution of online trust: Embedded in the SVM and NNT analyses, it was found that the emerged components of online trust changed over time independently. Ability that is associated with the cognitive evaluations played a greater role in the initial online trust building than benevolence and integrity did. On the contrary, benevolence that involved affective evaluations played a more important role than ability in the undermining of trust within the context of online forums. Results were in agreement with the findings by Dimoka (2010).

RQ 3: *How can the theory of critical mass be applied to understand the structural influence of online forums in relation to knowledge contribution continuance?*

- Investigative question five: how is the critical point beyond which the mass phenomenon of knowledge sharing within online forums achieved?
 - Investigative question six: what happens before and after the critical point in terms of the online knowledge contributions?
- i. Online forums are scale-free networks: a scale-free network is self-sustaining (Barabasi and Albert, 1999). Results of power-law fitting showed that the online forum in study three demonstrated the properties of scale-free networks. In addition, truncated power law rather than power-law was more suitable for describing scale-free networks, in agreement with the finding by Clauset *et al.* (2009).
 - ii. A phase transition within online forum in study three was observed and the critical point after which the online forum would be self-sustaining was quickly achieved: the network theories that were rarely incorporated into the analysis of web-based discussion groups should be applied in order to understand the theory of critical mass involved with phase transitions (e.g. Westland, 2010). Truncated power law is associated with phase transition in terms of its growth rate (e.g. Newman, 2003). Because the critical point over which the spinning cluster emerges was found to be quickly achieved, it could be understood that the concept of perceived critical mass would be suitable for explaining online knowledge sharing behaviours, in agreement with the argument by Cho (2011).
 - iii. After the critical point, it was the critical mass members within online forums who would ensure their continuing development: once the online forum under study has been developed into a scale-free network, the role of critical mass members in the evolution of the online forum was demonstrated through network attacking simulations using the characteristics of the empirical data. The network attacking method was embedded in the proposition by Albert *et al.* (2000).

Attacking only around 5% of critical mass members was sufficient to destroy the online forum in study three. However, it required random deletion of around 30% of members so that the online forum would be disconnected. Results were in agreement with the previous findings (e.g. Albert *et al.*, 2000). Moreover, the linear preferential attachment was found to be the mechanism that governs the development of online forums after the critical point (i.e. the connections to contributors are the linear proportion to their contributions), in agreement with Barabasi (2013) that the preferential attachment in the internet is measured around 1. It was concluded that an online forum could be sustainable as long as there were critical mass members who would pay the online contribution cost, which is consistent with the argument by Oliver and Marwell (1988). In return, critical mass members could be more popular (e.g. hubs can be more popular (Newman, 2005)), which might suggest their motivations to contribute online.

- iv. Before the critical point, it was the critical mass members who paid the start-up contribution cost: although results showed that the critical mass members were not necessarily duplicated during the evolution of online forums (i.e. new critical mass members could occur after the critical point), and the numbers of free-riders increased after the critical point (i.e. the connections to hubs tend to decrease in the late stage of its development), it was found that the critical mass members identified in the initial stage played an essential role in the membership expansion (i.e. the super-linear preferential attachment was found being a remarkable characteristic before the critical point). The finding of an increase numbers of free-riders along with the expansion in membership size after the critical point is in agreement with the previous empirical studies that the high connections items have a tendency to drop in connections (e.g. Blumm *et al.*, 2012; Centola, 2013). The finding of the linear preferential attachment in the final stage is in agreement with the previous studies that the attachment function measurement for internet base data is around 1 (e.g. Joeng *et al.*, 2001; Barabasi, 2013). However, few studies have sought to examine the changes in the network functions with the consecutive stages (e.g. Shmueli and Altshuler, 2014), study three further investigated what has happened before and after the critical point regarding to the

preferential attachment function and network structures in the context of online forums. Taken together, the return to the contributors in term of the connections associated to them was found to exceed their contribution cost. This is consistent with the proposition by Oliver and Marwell (1988) that public goods are value-added (i.e. the total values of public goods are greater than that by individual contributors), because the population is heterogeneous and some have more resources to contribute than others. However, public goods will never be created without the initial contributors (e.g. Ostrom, 2000), and it is those initial critical mass members who solve the start-up dilemma (e.g. Centola, 2013).

7.3 Theoretical implications

The study addressed the reasons for continuously contributing and sharing knowledge within online forums. The continuous contributions by members make the forum sustainable throughout time. This thesis is composed of three studies in order to answer the main research question with respect to the sustainability of knowledge contributions within online forums.

The contributions of study one are threefold. Firstly, the perspective taken in study one was comprehensive and considered various antecedents that are dynamic in nature. However, few studies have investigated the dynamic antecedents in the online knowledge continuance (e.g. Chen, 2007), and existing studies have used isolated or limited numbers of factors affecting on the online voluntary contribution behaviours embedded in different theoretical approaches (e.g. Zimmer *et al.*, 2010; Shen *et al.*, 2013). Knowledge sharing is an intentional behaviour (Ajzen, 1991). Grounded and combining insights from TPB (Ajzen 1991) and DTPB (Tylor and Todds 1995), study one developed the framework that investigated how the antecedents (namely trust in online forums, trust in members and perceived critical mass) interacted and influenced the determinants of intentional online contribution behaviours. The study considers that all these factors should be taken in consideration when analysing the intention to on-line sharing, and suggested that taking them individually only provides a partial perspective. For instance, it would be difficult to examine how these antecedents to the determinants of the online intentional

knowledge sharing behaviours impact together, and how they alter with each other within the nested models.

Secondly, previous studies have mainly investigated the causal relationships between the determinants (i.e. attitude, perceived behavioural control and subjective norms) and the response variable by developing theoretical hypotheses and examining them with SEM techniques (e.g. Chen, 2007; Shen *et al.*, 2013). The examinations on the identified antecedents of determinants have been little addressed previously, possibly as one consequence of the fact that the integrative approach has not been the main research stream with existing studies. For instance, limited studies have sought to investigate the relationships between interpersonal trust and normative beliefs (Jeffries and Becker, 2008). Findings from this study can add knowledge to the literature and show that the higher levels of benevolence/integrity based trust in members can result in higher levels of perceived normative pressures on members. In general, study one suggested that trust in online forums should be considered as a motivator for attitude and perceived behavioural control, and trust in members and perceived critical mass should be specified as antecedents for subjective norms.

Thirdly, how those antecedents impact together on the determinants of the intentional online contribution behaviours remains unknown. The moderated mediation examinations have contributed knowledge about the underlying relationships between antecedents, and this helps the understanding of how these causal predictors play a role in sustaining online forums in different magnitude levels. Again, this can be better examined within an integrative model, and has not been examined previously. No research has sought to test the causal relationships between interpersonal trust and perceived critical mass. Study one found out that trust in members who would like to share knowledge with others could be a predictor of the perceived size of contributors, but its effects were completely mediated by trust in online forums. Perceived critical mass partially mediated the path from trust in members regressing to subjective norms, but such mediation effects disappeared within the integrative model. Trust in members favoured trust in

online forums, in agreement with previous studies (e.g. Luo, 2006), and trust in online forums completely mediated the causal relationships between trust in members and attitude.

Study two has added to the literature on the dimensions of online trust (i.e. ability, benevolence and integrity, and predictability), and made a contribution to knowledge by identifying how individual dimensions change over time. Ability and benevolence were found to explain around 80% of variances explained in the overall trust within the context of online forums. Previous studies have generally treated online trust as unidimensional, and not explored the stability of dimensions over time (e.g. Palmer and Huo, 2013). The findings are in agreement with the study by Dimoka (2010) that, cognitive factors such as ability play a key role in the initial trust building, while, the undermining of online trust was more involved with the affective evaluations such as benevolence.

Study three, taking a dynamic view, has incorporated theories from network science to explore the role of critical mass members in sustaining online forums. Theories from network science are little integrated into studies in the field of social science (Barabasi, 2009). Few studies have sought to examine changes in the network functions with the consecutive stages (Shmueli and Altshuler, 2014). Results from study three showed that in the context of online forums, the super- linear preferential attachment was the remarkable function that governed the expansion of network in size before the critical point, and that the drop in the connections to the critical mass members was the other characteristic after the critical point. Although findings about what had happened after the critical point were in agreement with previous studies (e.g. Blumm *et al.*, 2012), only several studies (e.g. Centola, 2013) have sought to explain those findings using the established theories from social science. Previous studies on critical mass have mainly investigated the mass phenomena after the critical point (Cho, 2011). Study three has examined the online knowledge sharing phenomena before, near and after the critical point. Few studies have sought to investigate the role of the theory of critical mass applied in the context of Web 2 environments (e.g. Westland, 2010; Centola, 2013). Study three used an empirical study with a large scale of online forum

and incorporated the theory of critical mass to understand the knowledge continuance within online forums.

Taken together, this thesis proposed that strong and weak ties co-exist and play a role in online knowledge continuance. Although trust in online forums (involved with weak ties) has been found to have more weighted power in predicting the determinants for online intentional behaviours, and it completely mediated the effects of perceived critical mass regressed on trust in members, the role of strong ties cannot be ignored, which was demonstrated by the support of H5 that trust in members (involved with strong ties) can lead to trust in online forums. Results from study three led to a similar conclusion, because the strong ties created between the critical mass members can evoke the expansion in the size of memberships. It is worth noting that this finding is in agreement with the argument by Granovetter (1973) that weak ties are more likely to occur if the strong ties are presented. That is, the previous findings, generated with simulations (i.e. Centola, 2013), show that strong ties can impede the expansion of memberships, which were questioned with the empirical study. Additionally, the study has further clarified that the benevolence and integrity based trust in members were the antecedent of trust in online forums, because the former is associated with the perceived knowledge available within a forum that can reveal the ability of that forum, in agreement with the findings by Levin *et al.* (2003).

The findings of the study have added knowledge to the literature in four fundamental ways. Firstly, an integrative model should be considered in order to better understand how web-based discussion groups can be sustainable, because those antecedents are independent from each other. In addition, the tests of hypotheses have shown that paths from the social and structural antecedents to the intention to contribute online were all significant at 0.001 levels. Secondly, the casual relationships among antecedents are investigated, while existing studies provide little understanding about how these antecedents impact together on the online knowledge continuance. This is explained because previous studies tend to examine the antecedents within nested models, but have not sought to provide an integrative view. Thirdly, the dynamic aspects of the antecedents have been investigated, while the majority of previous

studies have taken the cross-sectional approach. Fourthly, the multiple dimensional concepts of online trust and critical mass have been examined in the context of online forums, which was little known in this field.

7.4 Managerial considerations

A numbers of previous studies, as well as the findings of this thesis, have shown the importance of online forums to the business and society (Vargo and Lusch, 2008; Wasco *et al.*, 2009). By developing an online forum, a firm can facilitate its customers to “co-create” value by sharing knowledge between firms and customers (Vargo and Lusch, 2008).

One particular managerial consideration with extended theoretical underpinnings was the program by online organizations that should be long-term oriented, because these antecedents to the determinants of online intentional contribution behaviours are dynamic in nature. For instance, trust in the context of online forums can be developed or undermined over time.

In addition, both the social and structural influences considerations should be incorporated into the policy. Online trust is associated with the social influences (e.g. Wasco and Faraj, 2005; Zimmer *et al.*, 2010). Critical mass that involves the percolation phenomenon and the influences of the perceived membership size on the intentional behaviours are considered as the structural factors (e.g. Westland, 2010). Study one has urged that the integrative models rather than the nested models should be considered in order to provide the complementary understandings in the online knowledge contribution behaviours, because the antecedents identified in the study have an influence on each other. For example, the partial mediation effects of the perceived critical mass on trust in members regressing to subjective norms can disappear within the integrative model proposed by study one.

Regarding to the social factors, findings from study one suggested that the competence and benevolence demonstrated by online forums have the weighted power in influencing online knowledge continuance. The findings from study two showed that the affective evaluations, such as an empathetic approach to

solving problems raised by customers are the key factors for trust building in brand in the later stage, which is in agreement with Dimoka (2010): that the undermining of trust is more associated with brain's emotional process. These results suggest that companies can develop their capability in terms of hosting numbers of members so that the critical mass members are encouraged to contribute knowledge within online forums, e.g. being more popular over time (Newman, 2005). For example, thousands or millions of members online simultaneously; limited moderation on the exchanges within members; voting systems that encourage members to publish quality content and so on.

Although the structural factor identified in this thesis, i.e. the concept of critical mass, is often ignored in previous survey – based studies, results from study three have shown that the theory of critical mass is important to understand the online knowledge continuance. The objective for managers is to have the firm-hosted online forums being scale-free, because this type of network is self-sustaining (e.g. Barabasi, 2013). Results from study three showed that a scale-free network is keeping on growing after the critical point, in agreement with the argument by Cohen *et al.* (2002).

One benefit for managers to develop the scale-free networks refers to the cost saving in the later stage, i.e. after the critical point. Findings from study three showed that only 5% of members who hold majority connections of an online forum are sufficient to ensure the ongoing development of online forums, but it required to remove around 30% of memberships to break down the functionality of self-sustaining. This is consistent with results generated from previous studies (Albert *et al.*, 2000). Moreover, findings from study three showed that the numbers of free-riders increased after the critical point, in agreement with the findings by Centola (2013). That is, strategies applied after the critical point only need to focus on a small group of critical mass members (5% of memberships in term of the connections) rather than satisfying everybody.

It is noted that the success of online discussion groups is not automatically associated with financial investments, but with the value of knowledge contributed by customers (e.g. Wasco *et al.*, 2009). This is particularly true in the initial stage, because this stage of an online forum's development involves random

factors (e.g. findings from study three showed that online networks were evolved from random networks, in agreement with the argument by Newman (2005)).

To solve the start-up problem so that the firm hosted online forums can be the scale-free networks, previous studies propose the use of incentives (e.g. Schulze *et al.*, 2014) in order to attract members to join in, but this can be expensive in particular for the small-medium size companies. An efficient method can lead to the reputation of a brand building through word-of-mouth during the implementation stages. Word-of-mouth communications refer to weak ties being created among members (e.g. Libai *et al.*, 2013). The findings from study one are consistent with the argument that weak ties play the role of 'bridge' that spans the communications among small groups involved with strong ties (e.g. Granovetter, 1973). The findings from study two suggested that the cognitive assessments such as the knowledge quality expected by members should be highlighted in the earlier trust building in online organisations (weak ties) (e.g. Haythornthwaite, 2002), in agreement with Domika (2010). The findings from study three indicated that the super linear preferential attachment was a remarkable characteristic that governed the communications among members before the critical point (i.e. the initial implementing stage). This referred to strong ties being created among the initial critical mass members which could promote wider communications involved with weak ties, in agreement with the findings by Hinz *et al.* (2011). All the three studies have showed that brand building through weak ties was essential during the initial stage.

In the context of online forums, brand building in the implementation stage is strongly associated with the knowledge quality that is attractive and memorable to the followers (e.g. Berger and Schwartz, 2011). Knowledge is the intrinsic reason why online forums exist (e.g. Wasco *et al.*, 2009). To attract experts in the initial stage, firms can provide awards / recognition for the early experts. Beside the online voting system, Wikipedia organises the annual meetings in different cities and invites the experts as the main speakers. If the financial conditions allow, firms are encouraged to gather experts and organise the off-line meetings that helps experts to have a sense of belonging to the online forums.

Additionally, the study argues that the critical mass members can be identified by the numbers of connections associated with them. This is because a relatively larger number of replies to the knowledge left by those critical mass members can reflect the value of knowledge contributed by them. The degree centrality measurement can be an efficient method, in particular for the small-and-medium size firms.

7.5 Limitations of the study

The study has offered a dynamic perspective embedded in a mixed methodology design to understand both the social and structural influences on sustaining online forums. The design has been seeking to interpret findings from different methods with the expansion intent. However, as a direct shortcoming of this methodology, the study has limitations in different ways which should be further considered:

- i. There is need for more detailed considerations, for instance of the paradigmatic issues associated with mixed methods. Bazeley (2002) argues that paradigmatic rose by mixed methods cannot be resolved, because one cannot prove paradigm. As a consequence of this, the interpretations of findings that generated from different studies may not be complete.
- ii. The study has explored the key antecedents of voluntary contribution online. However, other factor(s) were not considered. Indeed, the moderated mediation models to examine the mediation effect of general trust on interpersonal trust have suggested other potential antecedent(s) that might have impact on their relationship. Findings from the network structural analysis have suggested “reputation” as another factor that is one component of social capital as online trust could positively impact on motivational antecedents of online contributions.
- iii. Results generated from study two have provided open information. It was possible that the observed decline of online trust has been affected by other factors such as the lifecycle of brand. In addition, the empirical study was not designed to investigate either the association of cognitive evaluations with initial trust building or affective assessments impact on the decline of trust. Finally, although study two has taken a dynamic approach to collect data for a period of six years, individuals were not uniquely tracked during this period.

- iv. Due to a lack of experimental design and the large size of data, the growth mechanism of online forums in study three was not fully explored. In addition, other algorithms that can be used to identify the critical mass members were not tested.

7.6 Recommendations for future studies

Understanding the antecedents of sustaining online forums is therefore multifaceted. To generate achievable strategies and develop a relatively more complete framework with regard to sustainability of online knowledge sharing, there are needs for more empirical cases that are embedded in different worldviews to allow knowledge from different disciplines to be integrated. The following future studies can contribute to the fulfilment of this goal:

- i. Empirical studies that seek to understand the antecedents of continuous contribution behaviours online, and that are embedded in the dynamic perspectives. Although the importance to explore the dynamic antecedents of the determinants of online intentional knowledge sharing behaviours has been acknowledged, the majority of existing empirical studies embedded in the different theoretical frameworks are cross-sectional (e.g. Chen, 2007). Both study two and three have sought to feed this knowledge gap by taking a dynamic view. However, the analyses of these two studies are performed on the data representing two cases (skype and Stack overflow accordingly). As a result of which, more empirical studies in this field are necessary.
- ii. Studies with experimental designs in order to explore the complex causal relationships between social and structural influences on online contribution behaviours. To date, the understanding of the inter-relationships between the antecedents impacting on the online knowledge continuance is limited. The moderated mediation analyses performed in study one can add knowledge to the related literature. However, empirical studies with experimental designs using mediated moderation models can be further developed, in order to provide a more complete view on how these antecedents act together and influence the online knowledge sharing behaviours.

- iii. Studies that seek to evaluate online distrust, and its conceptualizations. Trust and distrust are two different and complex concepts (e.g. McKnight and Chervany, 2002). The results from study two can help to understand how online trust is conceptualized and evolves in the context of online forums. The future empirical studies can be taken to understand the evolution of online distrust that possibly affects online knowledge sharing behaviours.
- iv. Empirical studies that are designed to explore how online trust can be incorporated into network structural analysis. The results from study one show that online trust influence perceived critical mass in the context of online forums, such finding can inform, for instance, future studies on the prediction of trust that spreads or not over an online forum by using network analysis techniques.
- v. Empirical studies that test different diffusion-based algorithms to identify the important users within online forums. There are different algorithms that can be used to identify the influencers in a network (e.g. Ghoshal and Barabasi, 2011). The measurement of degree centrality is the mostly performed with existing studies. As a result of which, more studies that seek to compare the efficiency of different algorithms are required.
- vi. Studies that are embedded in experimental design to explore the phenomena of assortive and disassortive matches within online forums seen as networks in order to understand the role of those patterns in sustaining online forums. The results from study three show that experts are more likely to share knowledge between each other (assertive matches). Further developments for understanding the reasons why and how the phenomena of assertive or disassortive matches emerge are encouraged.

7.7 Summary conclusions

The study has investigated not only the social antecedents but also the structural influences on sustaining online knowledge sharing, which is consistent with argument by Ridings and Wasco (2010). Firstly an integrated model that encompasses the two sides of influences was examined using CB-SEM. Secondly the causal relationships between antecedents were examined by creating mediation and moderated mediation models. Thirdly results from study one informed the expansion stages within which the

dynamic aspects of antecedents have been explored using Webnography and network analysis approaches accordingly. The study has integrated knowledge from social and network science in order to expand upon findings in previous studies. In conclusion, an analysis of social and structural influences on the online voluntary knowledge contributions has found that both online trust and critical mass were the key dynamic antecedents, tested with statistical significance.

Appendix 1: Questionnaire

English version

Intention to share knowledge online

(Average completion time: 15 minutes)

Do you go to any online discussion group? May I invite you to answer the following questionnaires which are designed for a thesis? If you do agree on helping us:

Please complete all sections within the questionnaire. Try not to linger too long on any one section or question. Your first response will almost always be the best. This is not a test. There are no right or wrong answers. Please answer honestly. Your data remain confidential and are used for academic research only. The data are seen and analysed only by the academic researcher.

Section 1: This section asks your previous online experiences

1 Do you go to online forum/ online review website/ online community?

(If you choose —no, you don't need to continue.)

2. Would you like to list the name of your favourite online community/forum/discussion groups? (e.g. Ciao, skype, Youtube, StackOverflow , ect.)

Section 2: This section asks your intention to share information within online forum/community/discussion groups. Please choose one option that close to your answers.

Items	Strongly agree	Agree	Neutral	Disagree	Strongly Disagree
I try to share knowledge with online forums members. (INT1)					
I plan to share knowledge with online forums members. (INT2)					

I openly share information that I gained from news, magazines and journals with other online forums members. (INT3)					
I openly share my photo and camera related experiences or know-how with community members. (INT4)					
For me, sharing my knowledge with other members is pleasant. (ATT1)					
For me, sharing my knowledge with other members is enjoyable. (ATT2)					
For me, sharing my knowledge with other members is beneficial. (ATT3)					
For me, sharing my knowledge with other members is good. (ATT4)					
For me, sharing my knowledge with other members is valuable. (ATT5)					
It is always possible for me to share my knowledge with network members. (PBC1)					
If I want, I always could share knowledge with online forums members. (PBC2)					
I feel assured that technological structures are adequate at protecting me from any problems with information systems. (PBC3)					
I enjoy giving my true opinion, which is not risky.(PBC4)					
Members expend effort to maintain harmony in this forum. (SN1)					
There is a high level of cooperation (e.g. replying to other members' questions and comments) among members of the online forum. (SN2)					
Members are willing to sacrifice time and effort for the benefit of this online forum. (SN3)					
There is a high level of sharing among members of this online forum. (SN4)					
My forum is very competent. (TRCA1)					
My forum is able to satisfy its members.(TRCA2)					
I can expect good advices from my forum.(TRCA3)					
My forum is very concerned about the					

<p>ability of people to get along. (TRCB1)</p> <p>If a member required help, my forum's members would do their best to help.(TRCB2)</p> <p>My forum behaves in a consistence manner.(TRCI1)</p> <p>I feel fine using my forum's service since it generally fulfils its agreements.(TRCI2)</p> <p>My forum tried to be fair in dealings between members.(TRCI3)</p>					
<p>Members throw their hearts into the communities' affairs.(TRMI1)</p> <p>Members show that they all have good morals.(TRMI2)</p> <p>Member's suggestions are the best they can offer.(TRMI3)</p> <p>Members are very concerned about their ability to be friendly with each other.(TRMB1)</p> <p>Members will not deliberately interrupt during the course of a discussion.(TRMB2)</p> <p>Members will help each other solve problems.(TRMB3)</p> <p>Members have appreciated skills in relation to the topic we discuss.(TRMA1)</p> <p>Members have enough knowledge about the subject we discuss.(TRMA2)</p> <p>Members have specialized capacity that can add to our conversation.(TRMA3)</p>					
<p>Many people participate in the discussions.(PCMD1)</p> <p>Many of my friends participate in the discussions.(PCMD2)</p> <p>Many of my friends make comments.(PCMD3)</p> <p>I have friends who give valuable suggestions. (PCMLINK1)</p> <p>I know the member(s) who give valuable suggestions, and they become my online friend(s). (PCMLINK2)</p> <p>I have friends who are very active.</p>					

(PCMLINK3)					
I don't spend too much time on online discussions, but I enjoy knowledge provided by others.(PCMB1)					
Information from my forum exceeds my knowledge.(PCMB2)					
In my online forum, there are several members who give valuable suggestions because they have more resources to offer.(PCMBC1)					
In my online forum, there are always several members who give valuable suggestions.(PCMBC2)					
In my online forum, only several members are active, not many people make comments.(PCMGI1)					
If those active members quit my online forum, it will be a big loss.(PCMGI2)					

Section 3 This section asks several individual questions.

1. Please indicate your gender Male Female

2. Please indicate your age range

< 19 20—35 36 —45 > 45

3. Please indicate your education

Primary/middle school/High school

Bachelor level degree

Master level degree

Doctoral level degree

Other

Thank you for your completing this questionnaire.

Your participation in this survey is greatly appreciated!

Chinese version

请问您访问或者加入任何网络论坛或者网络社区吗？ 如果可以，能邀请您回答以下匿名的非商业用途的调查问卷吗？ 这里没有正确或者错误的答案，请根据您的第一感觉如实回答。

以下题目中出现的社区，请理解为您最喜欢的或者常访问的网络论坛，社交网络，群，等网络社区的同义词。

第一部分 敬请告之您的互联网经历

您访问或者加入网络论坛或者社区吗？ 如果您的回答为非，请放弃该问卷，谢谢您的参与。

如果您的回答为是，请问您喜欢访问的论坛或者社区的名字是什么呢？ (例如，CIAO，SKYPE，百度贴吧，新浪微博，天涯，QQ 群等)

第二部分 请回答下述关于参与 BBS 讨论意向相关问题。

问题	盛赞	善	不过尔尔	疑	非
我尽力与论坛或者社区的其他成员分享信息和知识。(INT1)					
我计划与论坛或者社区的其他成员分享信息和知识。(INT2)					
在我的社区，我公开分析新闻，杂志，报纸等内容和消息。(INT3)					
我会向社区成员公开自己的照片或者相关图片信息。(INT4)					
对我而言，与其他成员分析知识和信息是愉快的。(ATT1)					
对我而言，与其他成员分析知识和信息是享受的。(ATT2)					

对我而言，与其他成员分析知识和信息是有益的。(ATT3)					
对我而言，与其他成员分析知识和信息是好的。(ATT4)					
对我而言，与其他成员分析知识和信息是有用的。(ATT5)					
和其他成员分析知识总是可能的。(PBC1)					
如果我愿意，我总是可以和其他成员分享知识。(PBC2)					
我认为社区的技术过关，不会泄露我的个人信息，保护我的电脑受到病毒攻击。(PBC3)					
我喜欢给出自己真实的意见，这并未有风险。(PBC4)					
会员们尽力维持社区的和睦。(SN1)					
社区里的合作程度很高，例如，多回复，很少 0 回复。(SN2)					
会员们愿意花费时间精力参加社区的讨论或者活动。(SN3)					
会员们的分享信息的程度很高。(SN4)					
我的社区很有意义，有能力。(TRCA1)					
我的社区能够满足大伙儿的需求。(TRCA2)					
我可以从我的社区获得好的建议。(TRCA3)					
我的社区非常在意能够把活跃的或者有能力的人聚在一起。(例如 贴吧，QQ 召唤，@某某，邀请朋友，等) (TRCB1)					
如果社区里有会员寻求帮助，我的社区会尽力满足的。(TRCB2)					
我的社区管理员行为表现一致。					

<p>(TRCI1)</p> <p>我认为社区的服务是周到的，因为社区总的来说是履行了制定的条约的。(TRCI2)</p> <p>我的社区管理员尽力平等对待人和事。(TRCI3)</p>					
<p>社区的会员们很关心社区的活动或者情况。(TRMI1)</p> <p>会员们 道德观良好。(TRMI2)</p> <p>会员们给出的建议应该没有隐瞒的。(TRMI3)</p> <p>会员们不会故意歪贴，歪题。(TRMI1)</p> <p>会员们很在乎自己的态度是否友好。(TRMB1)</p> <p>社区的会员们通常有专业知识背景，能够发展讨论的话题，举一反三。(TRMA1)</p> <p>社区的会员们对讨论的话题有足够的背景知识。(TRMA2)</p> <p>社区的会员们可以对讨论的话题添加评论。(TRMA3)</p>					
<p>我认为很多人参加社区的讨论或者意见交流。(PCMD1)</p> <p>我有很多朋友参加社区的讨论或者意见交流。(PCMD2)</p> <p>我有很多朋友做评论。(PCMD3)</p> <p>社区里，我有朋友给出的建议或者意见很出众。(PCMLINK1)</p> <p>我认识出色的会员(们)，后来我们成为了朋友。(PCMLINK2)</p> <p>社区里，我有朋友非常活跃。(PCMLINK3)</p> <p>我并未花费很多时间和精力参加讨论活动，但是我很喜欢网友提</p>					

供的分享的信息或知识。 (PCMB1) 社区提供的知识信息丰富，我可以从中汲取营养。(PCMB2) 在我的社区，总是有少数会员能提出很好的见解和解决方案。 (PCMBC1) 在我的社区，总是有少数会员能提出很好的见解和解决方案，因为他们更专业些。(PCMBC2) 在我的社区，只是少数人非常活跃，大部分会员沉默。(PCMG1) 如果我的社区里那些尽管是少数的活跃会员离开了，这会给社区造成极大的损失。(PCMG2)					
--	--	--	--	--	--

第三部分 请回答以下 3 个关于个人信息的问题

1. 您的性别

F

M
2. 您的年龄

< 19

20—35

36 —45

> 45
3. 您的教育程度

小学/初中/高中

大学

硕士

博士

其他
- 非常感谢您的参与！您为我们提供了很大的帮助！
- Page291

L'intention de partager vos connaissances sur ligne

(Moyenne de remplir: 15 minuits)

Avisez-vous des expériences de partager vos connaissances sur ligne? Je me permets de vous adresser à remplir un questionnaire pour ma thèse. Je vous remercie en avance pour vos aidées.

Veuillez remplir tous les champs du questionnaire. Essayez-vous de ne pas prendre trop du temps en répondre les questions. Sachant qu'il ne s'agit pas d'un test, la peur de l'échec n'existe pas. L'important c'est de répondre très rapidement aux questions dans ce sondage. Il n'y aura pas de vrais ou faux réponse, et nous apprécions vos premières réflexions. Toutes vos réponses restent confidentielles et ne seront jamais loués ou vendus. Nous analyserons les données pour le but d'académie.

Première partie: Vos expériences sur ligne

1 Etiez-vous déjà allé sur un forum?

(Si vous répondez "non" à cette question, vous pourriez arrêter à remplir les champs suivants.)

2. Voulez-vous nous dire votre / vos forum(s) préféré(s)? (E.g. Ciao, skype, Youtube, StackOverflow , ect.)

Deuxième partie: Les questions suivantes vous demandent votre l'intention de partager vos connaissances sur le(s) forum(s). Veuillez cocher ce qui convient le plus.

Questions	Parfaitement	Correctement	Tant bien que mal	Désaccord	Désaccord parfaitement
J'essaye de partager mes connaissances avec les autres membres dans mon forum. (INT1)					
Je vais partager mes connaissances avec les autres membres dans mon forum. (INT2)					
Je partage des informations, des					

<p>journaux, des magazines, avec les autres membres dans mon forum. (INT3)</p> <p>Je partage mes photos et mes vidéos dans mon forum. ((INT4)</p>					
<p>Pour moi, c'est plaisante de partager mes savoir-faire dans mon forum. (ATT1)</p> <p>Pour moi, c'est agréable de partager mes savoir-faire dans mon forum. (ATT2)</p> <p>Pour moi, c'est bénéfique de partager mes savoir-faire dans mon forum. (ATT3)</p> <p>Pour moi, c'est bien de partager mes savoir-faire dans mon forum. (ATT4)</p> <p>Pour moi, c'est utile de partager mes savoir-faire dans mon forum. (ATT5)</p>					
<p>Il est toujours possible de partager mes connaissances avec les autres membres. (PBC1)</p> <p>Si je veux, je peux toujours partager mes connaissances avec les autres membres. (PBC2)</p> <p>Je suis assuré(e) que mon forum me protégera des problèmes, par exemple, des arnaques. (PBC3)</p> <p>Je suis ravi de donner mes vrais opinions, parce qu'il n'y a pas de risque dans mon forum, pour quoi que c'est soit. (PBC4)</p>					
<p>Les membres font leurs efforts pour maintenir des rapports harmonieux avec les autres. (SN1)</p> <p>Dans mon forum, le niveau de la coopération est haut. Par exemple, le taux de réponses aux questions posées par les membres est satisfait. (SN2)</p> <p>En général, les membres prennent leurs temps et font leurs efforts afin de contribuer ensemble pour les bénéfices de chacun(e). (SN3)</p> <p>Il est courant de partager des contenus dans mon forum. (SN4)</p>					
<p>Mon forum est compétent. (TRCA1)</p> <p>Mon forum est capable de satisfaire les besoins des membres. (TRCA2)</p> <p>Je pense que je pourrais avoir des bonnes idées depuis mon forum.</p>					

<p>(TRCA3) Mon forum essaie d'avoir des experts. (TRCB1)</p> <p>Les membres font leurs mieux pour répondre les besoins des autres. (TRCB2)</p> <p>Le management du forum est consistant. (TRCI1)</p> <p>Je profit les services de mon forum qui en général compromise leur promets. (TRCI2)</p> <p>Le management de mon forum est juste. (TRCI3)</p>					
<p>Les membres jettent leurs cœurs aux affaires de mon forum. (TRMI1)</p> <p>Les membres ont de la bonne humeur. (TRMI2)</p> <p>Les propositions des membres sont les meilleurs qu'ils peuvent offre. (TRMI3)</p> <p>Les membres cherchent à reste amicale avec les autres. (TRMB1)</p> <p>Les membres ne cherchent pas à interrompre des conversations en cours. (TRMB2)</p> <p>Les membres vont aider l'un pour l'autre. (TRMB3)</p> <p>Les membres peuvent apporter leurs expertises. (TRMA1)</p> <p>Les membres ont des savoir-faire. (TRMA2)</p> <p>Les membres sont des spécialistes. (TRMA3)</p>					
<p>Il y a du monde dans mon forum. (PCMD1)</p> <p>J'ai pas mal d'ami(e) dans mon forum. (PCMD2)</p> <p>Beaucoup de mes ami(e) laissent leurs commentaires dans mon forum. (PCMD3)</p> <p>J'ai des ami(e)s qui sont les experts dans mon forum. (PCMLINK1)</p> <p>Je connais des experts dans mon forum, et nous sont les ami(e)s maintenant. (PCMLINK2)</p>					

<p>J'ai des ami(e)s qui sont aussi les membres actifs de mon forum. (PCMLINK3)</p> <p>Je dépense peu de temps à discuter avec les autres dans mon forum. Cependant, j'apprécie leurs savoir-faire partagés. (PCMB1)</p> <p>J'apprends des autres membres dans mon forum. (PCMB2)</p> <p>Dans mon forum, il y a quelques membres qui donnent leurs avis précis parce qu'ils sont les experts dans les sujets en cours de discuter. (PCMBC1)</p> <p>Dans mon forum, il y a toujours quelques membres qui sont les experts sur les sujets en cours de discuter. (PCMBC2)</p> <p>Dans mon forum, quelques membres sont actifs, mais la plupart de gens sont silencieux. (PCMG1)</p> <p>Si quelques experts quittent mon forum, ce serait un choc énorme. (PCMG2)</p>					
--	--	--	--	--	--

Troisième partie: Nous nous permettons de vous demander quelques informations sur vous-même.

1. Vous êtes : Madame/Mademoiselle Monsieur

2. Vous avez

< 19ans 20—35ans 36 —45ans > 45 ans

3. Quel diplôme avez-vous:

L'école primaire/College/Licence Master 1 Master 2 Doctoral Autres

Nous vous remercions du temps que vous dédie à la réponse à cette enquête.

Votre collaboration est grandement appréciée!

Appendix 2: Examples of coding

Ability

	Examples	Score	Remarks
Competence, including service quality, system reliability	<p>1 Skype's service is undeniably brilliant, simple, convenient and cost-effective.</p> <p>2 The call quality from PC to PC is almost as clear as a telephone line.</p> <p>3 In terms of actual quality of the calls, well, they're above average.</p> <p>4 Skype really requires a broadband connection to work properly.</p> <p>5 The big problem with this service (and it IS a big problem) is that it is wholly unreliable.</p>	<p>5</p> <p>4</p> <p>3</p> <p>2</p> <p>1</p>	
Easy of use EOU	<p>1 The installation of Skype onto a Windows PC is incredibly easy.</p> <p>2 It is simple to download, and it really couldn't be easier.</p> <p>3 But to call on Skype the other person should also be on computer and logged into Skype.</p>	<p>5</p> <p>4</p> <p>3</p>	There are no score of "2" and "1" for EOU.
Perceived Usefulness PU	<p>1 I use it for both recreational and business and its great.</p> <p>2 With Skype out worldwide phonecall for cheap fares are possible - to landlines as well as to mobile phones.</p> <p>3 Overall for someone who makes international calls regularly I highly recommend skype, however for those who just want to make the occasional call every so often it may be better to just use a normal telephone line.</p>	<p>5</p> <p>4</p> <p>3</p>	This item is also included in "willingness to give a low price" with the score "4".

	Examples	Score	Remarks
	4 Rather than replace my landline as you might expect of a VOIP program, what it's done is replace a lot of instant messaging conversations with voice ones.	2	
	5 So - Skype might be okay for the free part of their service - but for serious business use and a proper customer care they are rubbish.	2	There is no score "1" for PU.
Perceived credibility	1 From all the services that are readily available out there I think this one is the best!	5	There are no score of "2" and "1" for perceive credibility.
	2 Skype has now probably become the world leader in terms of global communication over the internet.	4	
	3 In nowadays Internet world, Skype, like Ebay and Amazon, becomes a common word in daily life.	3	

Benevolence

	Example	Score	Remarks
Willingness to improve	1 Whenever there is a problem with the quality it usually for a very quick feedback. You can only transfer files/photos when both users are online.	4	There are no scores of "5" and "2" for "willingness to improve".
	2 Things are a bit better for Mac OS now, for a few weeks Skype 2.0 has been available for Mac's as well, which finally enables you to do video calls if you are not using MS WIndows.	3	This item is also included in "competence" with the score "3".
	3 But for me the issue which makes this product a firm NO! is the company's complete lack of willingness to deal with the security issues, and their actions towards worsening this.	1	This item is as well included in "ethical issues" with score "1".
Willingness to give a	1 Another reason I think Skype is	5	

	Example	Score	Remarks
low price or to give a promotion	great is because IT'S ALL FREE.		
	2 Keep your eyes open, sometimes Skype gives away free credit, as they did this autumn.	4	
	3 The rates for calls by SkypeOut are sounding really fair and as it runs with prepaid credits so you also can have a good overview over your costs and so will not getting shocked once to get a very high bill.	4	
	4 So as far as I can make out, it's not extravagantly expensive.	3	This item is considered as well in "Ethical issue" and "competence" with score "1" accordingly.
	5 This used to be free, but now to call phones you need skype credit.	2	
	6 Great in theory, but I am not a frequent user. Your Skype account will expire after 180 days of not using the pay as you go service. This means my skype credit expired twice by now and has cost me over £15! Their excuse is that accounting rules require it, that's daylight robbery.	1	
Handy or funny design	1 Landline numbers can also be incorporated into conference calls, making the system staggeringly flexible.	5	This item is also included in "competence" with the score "5".
	2 Overall this program is well designed and user friendly.	4	
	3 Overall, promising features, but improvement needed.	3	
	4 I'm not entirely sure that I like Skype's interface, though.	2	
	5 The user interface is scruffy.	1	

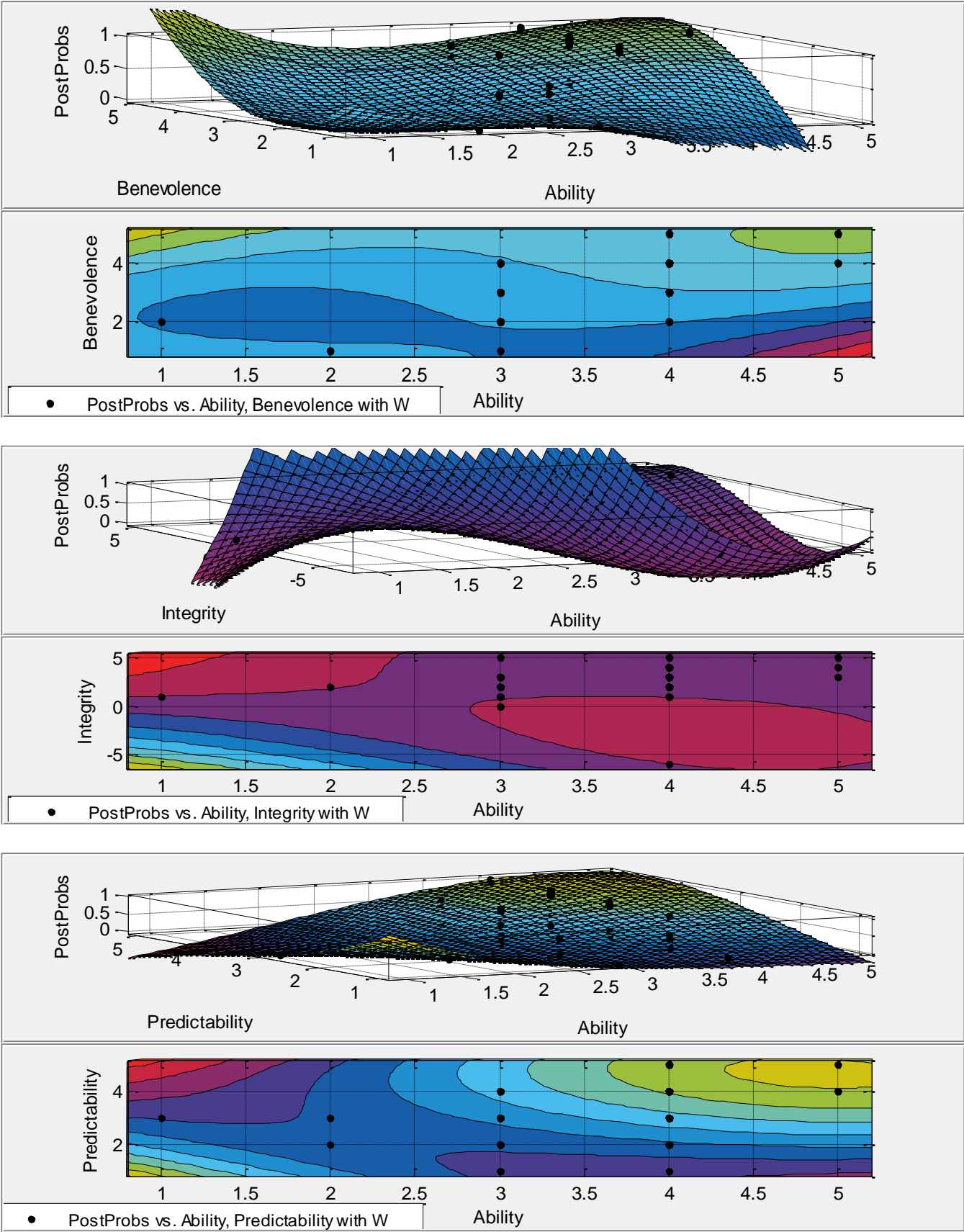
Integrity

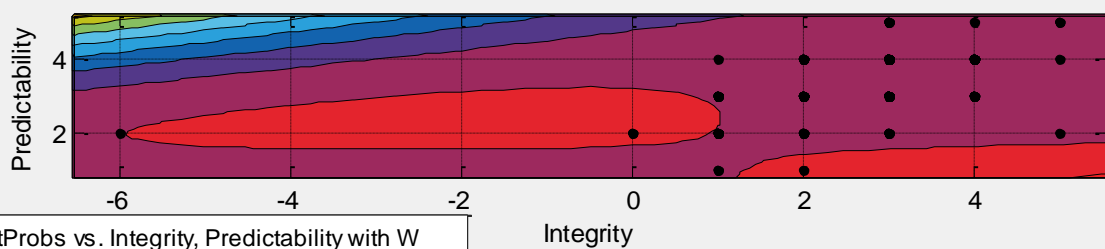
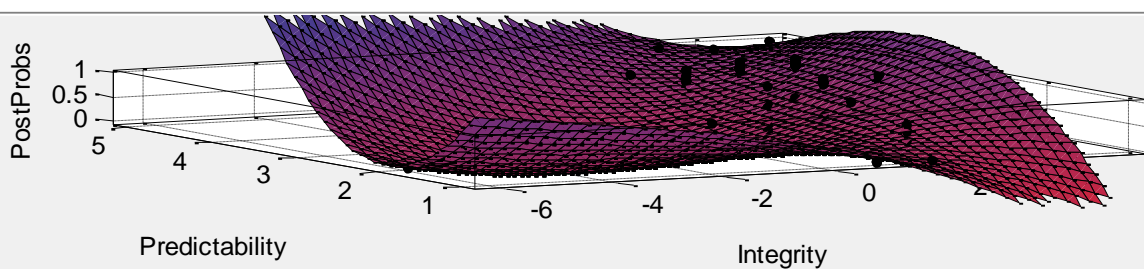
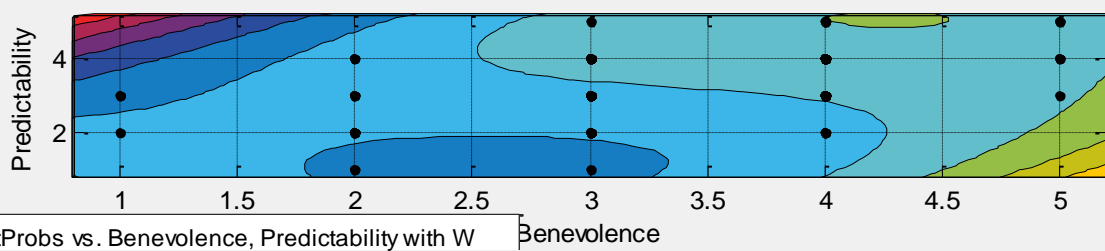
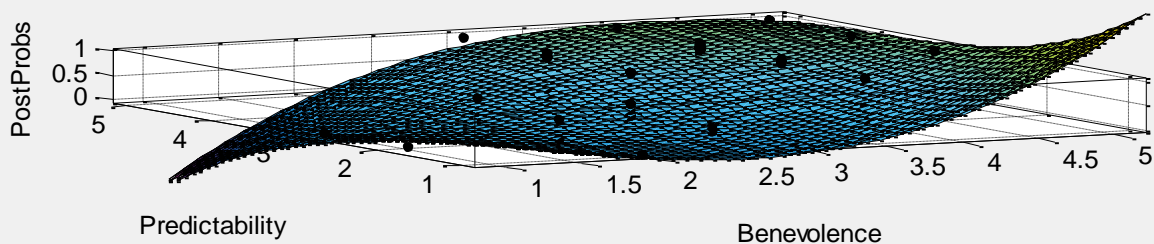
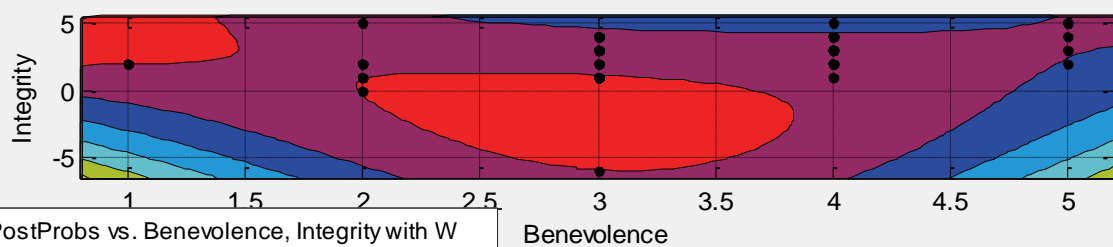
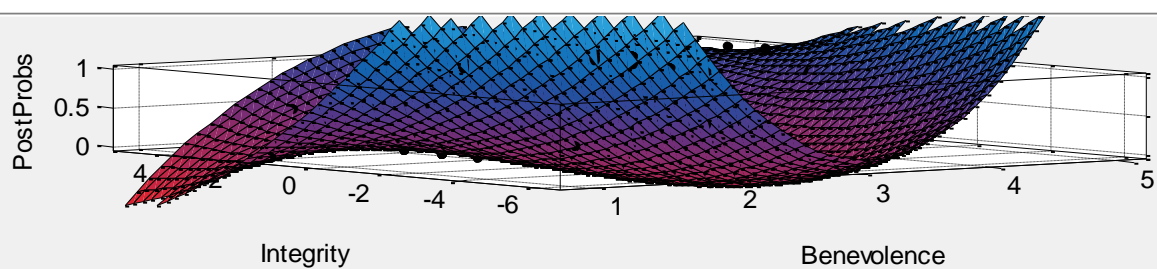
	Example	Score	Remarks
Expected outcomes	<p>1 This free software does exactly as the title suggests.</p> <p>2 After using it for work for a while and being fairly pleased with it's performance.</p> <p>3 On the whole could be improved but lets remember apart from the low start up cost of purchasing a Skype phone this service is FREE! and you dont get much nowadays for nothing.</p> <p>4 I had a SkypeIn number for almost 4 years. They had changed it along the way with little notice - that was bad - but I continued.</p> <p>5 I continued using Skype over the next 6-8 weeks, and have to say more often than not the line was terrible. Either they would hear me, but I couldn't hear them or vice versa. Sometimes calls would not connect, other times they would connect but drop out half way through the call. My frustration grew and I eventually gave up on this method too.</p>	<p>5</p> <p>4</p> <p>3</p> <p>2</p> <p>1</p>	<p>This item is also considered “willingness to improve” with the score “2”.</p>
Ethical Issues	<p>1 The page is 'padlocked' so you can be quite sure that your transaction will be safe.</p> <p>2 It is a good program if you like to talk but I suppose it can be a little bit exposed if young children are calling anyone that they want to.</p> <p>3 Many worry about others listening in on calls and to add to this worry, Skype will neither admit or deny this claim.</p> <p>4 Personally i do not trust the Skype system for sensitive information.</p>	<p>4</p> <p>3</p> <p>2</p> <p>1</p>	<p>There is no score “5” for “ethical issues”.</p>

Predictability

	Examples	Score	Remarks
Clear operation statement	1 For any problems that might appear Skype provides a detailed help and FAQ on their website, and all in English.	4	Clear statement helps users to predict what is going on in a near future.
	2 You can check the charges on the website. You can see mobile charges here as well, which are a little bit more expensive.	3	
Regularity	1 I'm charged roughly the same as a local call in the target country.	4	
	2 Overall Skype was good when they started but lately they had so many bugs in their service that there is a constant need to update all the time.	3	
	3 There is a POSSIBILITY that it will be updated, but it seems to be a very slim one as I saw no evidence of it in the time I used the application.	2	
Uncertainty	1 Skype records all call events including the missing calls. If there is a missing call, Skype will tell you in its Event.	4	
	2 Viruses are of course always a problem, but if you use Skype as it's intended to be used there is not much chance of getting one. But, as always, you have to be careful with receiving files. Spyware etc. are certainly not installed with Skype.	3	
	3 There is now a connection charge for SKYPE OUT calls, which is a nuisance if you get a lot of dropped calls.	1	

Appendix 3: posterior probabilities against paired predictors





Bibliography

- AIKEN, L. S., WEST, S. G., & RENO, R. R. (1991) Multiple regression: Testing and interpreting interactions. *Sage*.
- AJZEN, I. (1991) The theory of planned behavior. *Organizational Behavior and Human Decision Processes*. 50(2). p. 179–211.
- AJZEN, I. (2011) The theory of planned behaviour: reactions and reflections. *Psychology & health*. 26(9). p. 1113-1127.
- ALBERT, R., & BARABASI, A. L. (2002) Statistical mechanics of complex networks. *Reviews of modern physics*. 74(1). p. 47.
- ALBERT, R., JEONG, H., & BARABASI, A. L. (2000) Error and attack tolerance of complex networks. *Nature*. 406(6794). p. 378-382.
- ALGER, I. (2010) Public Goods Games, Altruism, and Evolution. *Journal of Public Economic Theory*. 12(4). p.789-813.
- ALLISON, P. (2014) Prediction vs. causation in regression analysis. *Statistical Horizons*. Available from: <http://statisticalhorizons.com/prediction-vs-causation-in-regression-analysis>
- ALLAMANIS, M., & SUTTON, C. (2013) Why, when, and what: analyzing stack overflow questions by topic, type, and code. In *Proceedings of the 10th Working Conference on Mining Software Repositories* (pp. 53-56). IEEE Press. Available from: <http://homepages.inf.ed.ac.uk/csutton/publications/msrCh2013.pdf>. [Accessed: 2013].
- ALSTOTT, J., BULLMORE, E., & PLENZ, D. (2014). powerlaw: a Python package for analysis of heavy-tailed distributions. *Plos one*. Available from: <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0085777>. [Accessed : 29th January 2014].
- AL-ALAWI, A. I., AL-MMARZOOQI, N. Y., & MOHAMMED, Y. F. (2007). Organizational culture and knowledge sharing: critical success factors. *Journal of knowledge management*. 11(2). p. 22-42.
- ANDERSEN, P.H. (2005) Relationship marketing and brand involvement of professionals through web-enhanced brand communities: The case of Coloplast. *Industrial Marketing Management*. 34(1). p.39-51.
- ANDERSON, J. C., & GERBING, D. W. (1988) Structural equation modeling in practice: A review and recommended two-step approach. *Psychological bulletin*. 103(3).p. 411.
- ANDREWS, J.C., DURVASULA, S. & AHHTER, S.H. (1990) A framework for Conceptualizing and measuring the involvement construct in advertising research. *Journal of Advertising*. 19(4). p. 27-40.
- ANNEN, K. (2003) Social capital, inclusive networks, and economic performance. *Journal of Economic Behavior & Organization*. 50(4). p.449-463.
- ARORA, R. (1982) Validation of an S-O-R model for situation, enduring, and response components of involvement. *Journal of Marketing Research*. 19(4). Special Issue on Causal Modeling. p. 505-516.

- ARMSTRONG, J. S., & OVERTON, T. S. (1977) Estimating nonresponse bias in mail surveys. *Journal of marketing research*. p. 396-402.
- AXELROD, R., & HAMILTON, W. D. (1981) The evolution of cooperation. *Science*. 211(4489).p. 1390-1396.
- BA, S. (2001) Establishing online trust through a community responsibility system. *Decision Support Systems*.31(3).p.323-336.
- BAEK, Y. M. & JUNG, C.S. (2015) Focusing the mediating role of institutional trust: How does interpersonal trust promote organizational commitment? *The Social Science Journal*. 52(4). p. 481-489.
- BAGOZZI, R. P. (2010) Consumer Intentions. *Wiley International Encyclopedia of Marketing*.
- BAGOZZI, R.P. & DHOLAKIA, U.M. (2002) Intentional social action in virtual communities. *Journal of Interactive Marketing*. 16(2). p.2-21.
- BAGOZZI, R.P. & DHOLAKIA, U.M. (2006) Antecedents and purchase consequences of customer participation in small group brand communities. *International Journal of Research in Marketing*. 23(1). p.45-61.
- BAGOZZI, R.P., YOUJAE, Y.& LYNN, W. P. (1991) Assessing construct validity in organizational research. *Administrative science quarterly*. p .421-458.
- BAK, P., TANG, C. & WIESENFELD, K. (1987) Self-organized criticality: An explanation of the 1/f noise. *Physical review letters*. 59(4). p .381.
- BAKER, E. L. (2003) Multiple measures: Toward tiered systems. *Educational Measurement: Issues and Practice*. 22(2). p . 13-17.
- BAMPO, M., MICHAEL, T. E., DINELI, R. M., DAVID, S . & MARK, W .(2008) The effects of the social structure of digital networks on viral marketing performance. *Information systems research*.19(3). p . 273-290.
- BARABASI, A. L. (2009) Scale-free networks: a decade and beyond. *Science*. 325(5939). p. 412.
- BARABASI, A. L. (2013) "Network science. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*. 371(1987). p. 20120375.
- BARABASI, A. L., & Albert, R. (1999) Emergence of scaling in random networks. *Science*. 286(5439). p. 509-512.
- BARABASI, A. L., DEZSO, Z., RAVASZ, E., YOOK, S. H., & OLTVAI, Z. (2003) Scale-free and hierarchical structures in complex networks. In *Modeling of Complex Systems: Seventh Granada Lectures* (661(1). p. 1-16). CA:AIP Publishing.
- BARABASI, A. L., & FRANGOS, J. (2014) *Linked: the new science of networks science of networks*. Basic Books.
- BARBER, B. (1983) The logic and limits of trust. *American Economic Review*. 64(1). p.1-10.

BARKI, H. & HARTWICK, J. (1994) Measuring user participation, user involvement, and user attitude. *Mis Quarterly*. 18(1). p.59-82.

BARON, R. M., & KENNY, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of personality and social psychology*.51(6). p.1173.

BART, Y., SHANKAR, V., SULTAN, F. & URBAN, G.L. (2005) Are the Drivers and role of online trust the same for all Web sites and consumers? A large-scale exploratory empirical study. *American Marketing Association*. 69(4).p.133-152.

BAUMGARTNER, H. & HOMBURG, C. (1996) Applications of structural equation modeling in marketing and consumer research: A review. *International journal of Research in Marketing*. 13(2). p. 139-161.

BAZELEY, P. (2002). *Issues in Mixing Qualitative and Quantitative Approaches to Research*. Available from : <http://www.researchsupport.com.au/MMIssues.pdf>. [Acceded 10th April 2002].

BEARDEN, W.O. & ETZEL, M.J. (1982) Reference group influence on product and brand purchase decisions. *Journal of Consumer Research*.9 (2). p. 183-194.

BECKER,G.S. (1974) A Theory of Social Interactions. *Journal of Political Economy*. 82(6). p. 1063-1093.

BEN-NER, A. & PUTTERMAN, L. (2000) On some implications of evolutionary psychology for the study of preferences and institutions. *Journal of Economic Behavior & Organization*. 43(1).p.91-99.

BENTLER, P. M. (1983) Some contributions to efficient statistics in structural models: Specification and estimation of moment structures. *Psychometrika*. 48(4). p. 493-517.

BENTLER,P. M. (1990) Comparative fit indexes in structural models. *Psychological bulletin*. 107(2). p. 238.

BENTLER, P. M. (2007) On tests and indices for evaluating structural models. *Personality and Individual Difference*. 42(5). p. 825-829.

BENTLER, P. M., & RAYKOV, T. (2000) On measures of explained variance in nonrecursive structural equation models. *Journal of Applied Psychology*. 85(1). p. 125.

BERGER, J., & SCHWARTZ, E. M. (2011) What drives immediate and ongoing word of mouth?. *Journal of Marketing Research*. 48(5). p. 869-880.

BHAGAT, P.S., KLEIN, A. & SHARMA, V. (2009) The impact of new media on internet-based group consumer behaviour. *Journal of Academy of Business and Economics*. 9 (3). p.83-94.

BICCIERI, C., XIAO, E., & MULDOON, R. (2011) Trustworthiness is a social norm, but trusting is not. *Politics, Philosophy & Economics*. 10(2). p. 170-187.

BICCIERI,C. & MULDOON, R. (2014) *Subjective norms*. The Stanford Encyclopedia of Philosophy (Spring 2014 Edition), Edward N. Zalta (ed.), Available from:
<<http://plato.stanford.edu/archives/spr2014/entries/social-norms/>>.

BISHOP, J. (2008) Enhancing the understanding of genres of web-based communities: The role of the ecological cognition framework. *International Journal of Web Based Communities*. 5(1). p. 4-17.

BJØRNSKOV, C. (2006) The multiple facets of social capital. *European Journal of Political Economy*. 22 (1). p. 22 –40.

BLAIKIE, N. W. H. (2000) Designing social research : the logic of anticipation. Cambridge. UK : Malden, MA: Polity Press : Blackwell.

BLUM, L., CUCKER, F., SHUB, M. & SMALE, S. (2012) Complexity and real computation. *Springer Science & Business Media*.

BOCK, G. W., ZMUD, R. W., KIM, Y. G. & LEE, J. N. (2005) Behavioral intention formation in knowledge sharing: Examining the roles of extrinsic motivators, social-psychological forces, and organizational climate. *MIS quarterly*. p. 87-111.

BOLLEN, K. A. (1989) A new incremental fit index for general structural equation models. *Sociological Methods & Research*. 17(3). p. 303-316.

BOLLOBAS, B., RIORDAN, O. ,SPENCER, J. & TUSNADY, G. (2001) The degree sequence of a scale-free random graph process. *Random Structures and Algorithms*. (18). p. 279-290.

BORGATTI, S.P. & EVERETT, M.G.(1992) Notions of position in social network analysis. *Sociological Methodology*. 22(1).p.1-35.

BONNEMAIZON, A., COVA, B. & LOUYOT, M.C. (2006) Relationship Marketing in 2015: A Delphi approach. *European Management Journal*. 25(1). p. 50-59.

BOO, S., BUSSE, J. & BALOGLU, S. (2009) A model of customer-based brand equity and its application to multiple destinations. *Tourism Management*.30(2).p. 219-231.

BOUMA, J., BULTE, E. & SOEST, D.V. (2008) Trust and cooperation: Social capital and community resource management. *Journal of Environmental Economics and Management*.56(2) .p.155-166.

BOULIANNE, E. & CHO, C.H. (2009) The rise and fall of WebTrust. *International Journal of Accounting Information Systems*. 10(4). p.229-244.

BOYD, J. (2002) In community we trust: Online security communication at eBay. *Journal of Computer-Mediated Communication*. 7(3).p.0-0. Available from: In community we trust: Online security communication at eBay. [Accessed: 23th June 2006].

BRANDON, R. (1994). *Making it Explicit: Reasoning, Representing, and Discursive Commitment*. Cambridge, MA: Harvard University Press.

BRAF, E. (2002) Knowledge or Information. *In Organizational Semiotics*. Springer US. p. 71-90.

BRATMAN, M. (1984) Two faces of intention. *The Philosophical Review*. 93(3). p. 375-405.

BREAZEALE, M. (2008) Word of mouth- an assessment of electronic word-of-mouth research. *International Journal of Market Research*. 51(3). p. 297-318.

BRENGMAN, M., & KARIMOV, F. P. (2012) The effect of web communities on consumers' initial trust in B2C e-commerce websites. *Management Research Review*.35(9). p. 791-817.

BRODIE, R.J., WHITTOME, J.R.M. & BRUSH, G.J. (2009) Investigating the service brand: A customer value perspective. *Journal of Business Research*. 62(3). p.345-355.

BROWN, J. D. (2009) Choosing the right type of rotation in PCA and EFA. *JALT Testing & Evaluation SIG Newsletter*. 13(3). p.20-25.

BROWN, J., BRODERICK, A.J. & LEE, N. (2007) Word of mouth communication within online communities: Conceptualising the online social network. *Journal of Interactive Marketing*. 21(3). p.2-20.

BROWNE, M. W. (2001) An overview of analytic rotation in exploratory factor analysis. *Multivariate Behavioral Research*. 36(1). p.111-150.

BRUNK, K.H. (2010) Exploring origins of ethical company/brand perceptions – A consumer perspective of corporate ethics. *Journal of Business Research*. 63(3). p.255- 262.

BRÜTTNER, O. B. & GÖRITZ, A. S. (2008). Perceived trustworthiness of online shops. *Journal of Consumer Behaviour*. 7(1). p. 35-50.

BYRNE, B.M. (2013) Structural equation modeling with AMOS: Basic concepts, applications, and programming. *Routledge*.

BORSBOOM, D., MELLENBERGH, G.J.,VAN HEERDEN, J. (2003) The Theoretical Status of Latent Variable. *Psychological Review*. 110 (2). p. 203–219. doi:10.1037/0033-295X.110.2.203

BUCHANAN, J.M. (1966) Joint supply, externality and optimality. *Economica*. 33(132). p. 404-415.

BUGHIN, J., DOOGAN, J. & VETVIK, O.J. (2010) A new way to measure word-of mouth marketing. *Mckinsey Quarterly*. 2.p.113-116.

BULLER, D.J. (2005) Evolutionary psychology: the emperor's new paradigm. *Trends in Cognitive Science*. 9(6). p. 277-283.

BUTLER, J.K. (1999) Trust expectations, information sharing, climate of trust, and negotiation effectiveness and efficiency. *Group & Organization Management*. 24(2). p.217-238.

BUTTNER, O.B. & GORITZ, A.S. (2008) Perceived trustworthiness of online shops. *Journal of Consumer Behavior*. 7(1). p.35-50.

BURNKRANT, R.E. & COUSINEAU, A. (1975) Information and Normative Social Influence in Buyer Behaviour. *Journal of Consumer Research*. 2(3). p.206-215.

BURTON, R. M. & KRANE, M. (1989) Density and uniqueness in percolation. *Communications in mathematical physics*, 121(3). p.501-505.

BYRNE, B. M. (ed.)(2013) *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. Routledge.

- CALDARELLI, G., CAPOCCI, A., DE LOS RIOS, P., & MUNOZ, M. A. (2002) Scale-free networks from varying vertex intrinsic fitness. *Physical review letters*. 89(25). p.258702.
- CARAGEA, D., SILVESCU, A. & HONAVAR, V. (2001) Analysis and synthesis of agents that learn from distributed dynamic data sources. In Emergent neural computational architectures based on neuroscience (p. 547-559). Springer Berlin Heidelberg.
- CARLSON, B.D., SUTER, T.A. & BROWN, T.J. (2008) Social versus psychological brand community: The role of psychological sense of brand community. *Journal of Business Research*. 61(4).p.284-291.
- CARUANA, A. & EWING, M.T. (2009) How corporate reputation, quality, and value influence online loyalty. *Journal of Business Research*. 63(9).p.1103-1110.
- CASALÓ, L.V., FLAVIAN, C. & GUINALIU, M. (2010) Relationship quality, community promotion and brand loyalty in virtual communities: Evidence from free software communities. *International Journal of Information Management*. 30(4). p. 357-367.
- CASTELEYN, J., MOTTART, A. & RUTTEN, K. (2008) How to use Facebook in your market research. *International Journal of Market Research*. 51(4). p.439-447.
- CENTOLA, D. (2013). A simple model of stability in critical mass dynamics. *Journal of statistical physics*, 151(1-2), p .238-253.
- CENTOLA, D. (2015). The social Origins of Networks and Diffusion. *American Journal of Sociology*. 120(5). p . 1295-1338.
- CHAN, K.W. & LI, S.Y. (2008) Understanding consumer-to-consumer interactions in virtual communities: The salience of reciprocity. *Journal of Business Research*. 63(9). p.1033-1040.
- CHANG, M. K., CHEUNG, W., & LAI, V. S. (2005) Literature derived reference models for the adoption of online shopping. *Information & Management*. 42(4). p. 543-559.
- CHANG, S.A. & TAYI, G. K. (2009) An analytical approach to bundling in the presence of customer transition effects. *Decision Support Systems*.48(1). p. 122-132.
- CHARKI, M.H. & JOSSE RAND, E. (2008) Online reverse auctions and the dynamics of trust. *Journal of Management Information Systems*. 24(4).p.175-197.
- CHATZIDAKIS, A. & MITUSSIS, D. (2007) Computer ethics and consumer ethics: The impact of the internet on consumers' ethical decision-making process. *Journal of Consumer Behaviour*. 6(5).p.305-320.
- CHENNAMANENI, A. (2007) Determinants of knowledge sharing behaviors: Developing and testing an integrated theoretical model. Available from : <https://uta-ir.tdl.org/uta-ir/handle/10106/305>
- CHIN, A. J., WAFI, S. A. W. S. K., & OOI, A. Y. (2009) The effect of internet trust and social influence towards willingness to purchase online in Labuan, Malaysia. *International Business Research*.2(2).p.72.
- CHIU, C. M., HSU, M. H., & WANG, E. T. (2006). Understanding knowledge sharing in virtual communities: An integration of social capital and social cognitive theories. *Decision support systems*, 42(3). p. 1872-1888.

- CHEN, I. Y. (2007). The factors influencing members' continuance intentions in professional virtual communities-a longitudinal study. *Journal of Information Science*. 33(4). p.451.
- CHENNAMANENI, A., TENG, J. T., & RAJA, M. K. (2012) A unified model of knowledge sharing behaviours: theoretical development and empirical test. *Behaviour & Information Technology*. 31(11). p.1097-1115.
- CHEUNG, C.M.K. & LEE, M.K.O. (2009) Understanding the sustainability of a virtual community: model development and empirical test. *Journal of Information Science*. 35(3). p.279-298.
- CHIU, C. M., CHANG, C. C., CHENG, H. L., & FANG, Y. H. (2009) Determinants of customer repurchase intention in online shopping. *Online information review*. 33(4). p. 761-784.
- CHIU, C. M., HSU, M. H., & WANG, E. T. (2006) Understanding knowledge sharing in virtual communities: an integration of social capital and social cognitive theories. *Decision support systems*. 42(3). p.1872-1888.
- CHRISTOPHER, S., WATTS, V., & YONG, S. (2008) Building and maintaining trust in a community-based participatory research partnership. *American Journal of Public Health*. 98(8). p.1398-1406.
- CHO, H. (2011) Theoretical intersections among social influences, beliefs, and intentions in the context of 3G mobile services in Singapore: Decomposing perceived critical mass and subjective norms. *Journal of Communication*. 61(2). p. 283-306.
- CHO, J. (2006) The mechanism of trust and distrust formation and their relational outcomes. *Journal of Retailing*. 82(1). p.25-35
- CHRISTENSEN, K. (2002) Percolation theory. *Imperial College London, London*, 40.
- CHU, S.C., LEUNG, L.C., HUI, Y.V. & CHEUNG, W. (2007) Evolution of e-commerce Web-sites: A conceptual framework and a longitudinal study. *Information & Management*. 44(2). p. 154-164.
- CURCHILL Jr, G. A. (1979) A paradigm for developing better measures of marketing constructs. *Journal of marketing research*. p. 64-73.
- CHURCHILL Jr, G. A., & IACOBucci, D. (2005) Marketing Research: Methodological. *Foundations (California: Sage)*.
- CLAUSET, A., SHALIZI, C. R. & NEWMAN, M. E. (2009) Power-law distributions in empirical data. *SIAM review*. 51(4). p.661-703.
- CLEMONS, E.K. (2009) The complex problem of monetizing virtual electronic social networks. *Decision Support Systems*. 48(1) .p.46-56.
- CLEMONS, E. K., BARNETT, S., & APPADURAI, A. (2007) The future of advertising and the value of social network websites: some preliminary examinations. *International conference on electronic commerce*. New York : Minneapolis, MN, ACM Press.
- COHEN, J. (1988) Statistical power analysis for the behavioral sciences(2nd ed.). Hillsdale, NJ: Erlbaum
- COHEN, R., BEN-AVRAHAM, D. & HAVLIN, S. (2002) Percolation critical exponents in scale-free networks. *Physical Review E*. 66(3). p. 036113.

- COHEN, R., EREZ, K., BEN-AVRAHAM, D., & HAVLIN, S. (2001) Breakdown of the Internet under intentional attack. *Physical review letters*. 86(16). p. 3682.
- COHEN, R. & HAVLIN, S. (2003) Scale-free networks are ultrasmall. *Physical review letters*.90(5). p.058701.
- COHEN, P. R., & LEVESQUE, H. J. (1990) Intention is choice with commitment. *Artificial intelligence*. 42(2-3). p. 213-261.
- COHEN, R., ROZENFELD, A. F., SCHWARTZ, N., BEN-AVRAHMA, D., & HAVLIN, S. (2003) Directed and non-directed scale-free networks. In *Statistical Mechanics of Complex Networks* (p. 23-45). Berlin Heidelberg: Springer.
- COLEMAN, J. S. (1988) Social capital in the creation of human capital. *American journal of sociology*. p. S95-S120.
- COMLEY, P. (2008) Online research communities. *International Journal of Market Research*. 50(5). Web 2.0 Special Issue. p.679-694.
- CONRAD, K. (2013) Probability distributions and maximum entropy. *retrieved November*. p .14.
- COOPER, D.R. & SCHINDLER, P.S. (ed.) (2008) *Business Research Methods*. London, UK: McGraw-Hill .
- CORTES, C. & VAPNIK, V. (1995) Support-vector networks. *Machine learning*. 20(3). p .273-297.
- CORTINA, J. M. (1993) What is coefficient alpha? An examination of theory and applications. *Journal of applied psychology*. 78(1). p . 98.
- COX, D.F. (1967) Risk handling in consumer behaviour- An intensive study of two cases. In D.F. Cox (eds.) *Risk Taking and Information Handling in Consumer Behavior*. Boston: Harvard University.
- COX, D.F. & RICH, S.U (1964) Perceived risk and consumer decision-marking: The case of telephone shopping. *Journal of Marketing Research*.1(4). p.32-39.
- CREESE, J. (2007) Web 2.0/Business 2.0: New web technologies. *organisations and WCM*.
- CROFT, M. (2008) Consumers in control. *Marketing Week*. 31(14). p. 29-30.
- CRONBACH, L. J. (1951) Coefficient alpha and the internal structure of tests. *Psychometrika*. 16(3). p. 297-334.
- CROSSLEY, N. & IBRAHIM, J. (2012) Critical mass, social networks and collective action: Exploring student political worlds. *Sociology*. 0038038511425560.
- CUMMINGS, J. A., BONK, C. J., & JACOBS, F. R. (2002) Twenty-first century college syllabi: Options for online communication and interactivity. *The Internet and Higher Education*. 5(1). p.1-19.
- CUMMINGS, J. N., BUTLER, B., & KRAUT, R. (2002) The quality of online social relationships. *Communications of the ACM*. 45(7). p.103-108.

- DARLEY, J. M. & LATANE, B. (1968) Bystander intervention in emergencies: diffusion of responsibility. *Journal of Personality and Social Psychology*. 8. p. 377-383.
- DAS, T. K. & TENG, B. S. (2004) The risk-based view of trust: A conceptual framework. *Journal of Business and Psychology*. 19(1). p. 85-116.
- DAVIS, F.D. (1993) User acceptance of information technology: System characteristics, user perceptions, and behavioral impact. *International Journal of Man-Machine Studies*. 38(3). p. 475-487.
- DAVIS, F.D., BAGOZZI, R.P. & WARSHAW, P.R. (1989) User acceptance of computer technology: a comparison of two theoretical models. *Management Science*. 35(8). p. 982-1003.
- DAWES, R. M. (1980) Social dilemmas. *Annual Review of Psychology*. 31(1). p. 169-193.
- DAWES, J. G. (2008) Do data characteristics change according to the number of scale points used? An experiment using 5 point, 7 point and 10 point scales. *International journal of market research*. p. 51(1).
- DAWKINS, R. (1989) The selfish gene. 1976. revised edn. UK: Oxford.
- DEHAENE, S., IZARD, V., SPELKE, E., & PICA, P. (2008) Log or linear? Distinct intuitions of the number scale in Western and Amazonian indigene cultures. *Science*. 320(5880). p. 1217-1220.
- DE KERVENOAEL, R., AYKAC, D.S.O. & PALMER, M. (2009) Online social capital: Understanding e-impulse buying in practice. *Journal of Retailing and Consumer Services*. 16(4). p. 320-328.
- DE VALCK, K., VANBRUGGEN, G.H. & WIERENGA, B. (2009) Virtual communities: A marketing perspective. *Decision Support Systems*. 47(3). p. 185-203.
- DELONE, W.H.& MCLEAN, E.R. (1992) Information systems success: the quest for the dependent variable. *Information Systems Research*. 3(1). p. 60-95.
- DELONE, W.H.& MCLEAN, E.R. (2003) The Delone and McLean model of information systems success: A ten-year update. *Journal of Management Information Systems*. 19(4). p.9-30.
- DEMANGE, G. (2010) Sharing information in Web communities. *Games and Economic Behavior*. 68(2).p. 580-601.
- DE MOOR, A. & WEIGAND, H. (2007) Formulising the evolution of virtual communities. *Information Systems*. 32(2). p. 223-247.
- DENNIS, C., MERRILEES, B., JAYAWARDHENA, C. & WRIGHT, L.T. (2009) E-consumer behaviour. *European Journal of Marketing*. 43(9/10). p. 1121-1139.
- DEQUECH, D. (2009) Institutions, subjective norms, and decision-theoretic norms. *Journal of Economic Behavior & Organization*. 72(1). p. 70-78.
- DEUTSCH, M. (1960a) The effect of motivational orientation upon trust and suspicion. *Human Relations*. 13. p. 123-139.
- DEUTSCH, M. (1958) Trust and suspicion. *Journal of Conflict Resolution*. 2. p. 264-279.

- DEUTSKENS, E., DE RUYTER, K., WETZELS, M. & OOSTERVELD, P. (2004) Response rate and response quality of internet-based surveys: An experimental study. *Marketing letter*. 15(1). p. 21-36.
- DE VALCK, K., VANBRUGGEN, G.H. & WIERENGA, B. (2009) Virtual communities: A marketing perspective. *Decision Support Systems*. 47(3). p. 185-203.
- DEVARAJ, S., FAN, M. & KOHLI, R. (2006) Examination of online channel preference: Using the structure-conduct-outcome framework. *Decision Support Systems*. 42(2). p. 1089-1103.
- DHOLAKIA, U.M. (1997) An investigation of the relationship between perceived risk and product involvement. *Advance in Consumer Research*. 24.p.159-167.
- DHOLAKIA, U.M. (2001) A motivational process model of product involvement and consumer risk perception. *European Journal of Marketing*. 35(11/12). p.1340-1360.
- DHOLAKIA, U.M., BAGOZZI, R.P. & PEARO, L.K. (2004) A Social influence model of consumer participation in network-and small-group-based virtual communities. *International Journal of Research in Marketing*. 21(3). p. 241-263.
- DHOLAKIA, U.M., BLAZEVIC, V., WIERTZ, C. & ALGESHEIMER, R. (2009) How customers benefit from participation in firm-hosted virtual P3 communities. *Journal of Service Research*. 12(2). p. 208-226.
- DIAMONTOPOLOUS, A., SIGUAW, J. A. & SIGUAW, J. A. (2000) Introducing LISREL: A guide for the uninitiated. *Sage*.
- DIAMANTOPOLOUS, A., & WINKLHOFER, H. M. (2001) Index construction with formative indicators: An alternative to scale development. *Journal of marketing research*. 38(2). p. 269-277.
- DIESENDORF, M. (2000) Sustainability and sustainable development. *Sustainability: The corporate challenge of the 21st century*. 2. p. 19-37.
- DIESTEL, R. (2005) Graph Theory (3rd ed.). Springer. ISBN 3-540-26182-6.
- DUFWENBERG, M., GACHTER, S & HENNIG-SCHMIDT, H. (2011) The Framing of Games and the Psychology of Play. *Games and Economic Behavior*. 23(2). p. 199-216.
- DIMOKA, A. (2010) What does the brain tell us about trust and distrust? Evidence from a functional neuroimaging study. *Mis Quarterly*. 34(2). p.373-396.
- DHOLAKIA, U. M., BAGOZZI, R. P., & PEARO, L. K. (2004) A social influence model of consumer participation in network-and small-group-based virtual communities. *International journal of research in marketing*. 21(3). p. 241-263.
- DODDS, P. S., & WATTS, D. J. (2004) Universal behavior in a generalized model of contagion. *Physical review letters*. 92(21). p.218701.
- DONALDSON, L. (2001) The contingency theory of organizations. *Sage*.
- DONALDSON, T. & DUNFEE, T.W. (2002) Ties that bind in business ethics: Social contracts and why they matter. *Journal of Banking & Finance*. 26(9). p. 1835-1865.

DOROGOVTSSEV, S. N., MENDES, J. F. F. & SAMUKHIN, A. N. (2001) Size-dependent degree distribution of a scale-free growing network. *Physical Review E*.63(6). p. 062101.

DYER, J. H., & CHU, W. (2003) The role of trustworthiness in reducing transaction costs and improving performance: Empirical evidence from the United States, Japan, and Korea. *Organization science*. 14(1). p. 57-68.

DYER, J. & NOBEOKA, K. (2002) Creating and managing a high performance knowledge-sharing network: the Toyota case. Available from: <http://dspace.mit.edu/handle/1721.1/1441>. [Accessed: 10th July 2002].

EASTLICK, M. A., LOTZ, S.L. & WARRINGTON, P. (2006) Understanding online B-TO-C relationship: An integrated model of privacy concerns, trust and commitment. *Journal of Business Research*.59(8). p.877-886

ECCLESTON, D. & GRISERI, L. (2008) How does Web 2.0 stretch traditional influencing patterns?. *International Journal of Market Research*. 50 (5). Web 2.0 Special Issue.

EASLEY, D. & KLEINBERG, J. (2010) *Networks, crowds, and markets: Reasoning about a highly connected world*. UK: Cambridge University Press.

EDWARDS, J. R. & LAMBERT, L. S. (2007) Methods for integrating moderation and mediation: a general analytical framework using moderated path analysis. *Psychological methods*. 12(1). p.1.

ERDEN, Z., VON KROGH, G. & KIM, S. (2012) Knowledge Sharing in an Online Community of Volunteers: The Role of Community Munificence. *European Management Review*. 9(4). p. 213-227.

ENDERS, A., HUNGENBERG, H., DENKER, H.P. & MAUCH, S. (2008) The long tail of social networking: Revenue models of social networking sites. *European Management Journal*. 26(3). p.199-211.

ERDÖS, P. & RENYI, A. (1960) On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*. 5(17-61). p. 43.

ERMISCH, J., GAMBETTA, D., LAURIE, H., SIEDLER, T. & UHRIG, S.C. N. (2009) Measuring people's trust. *Royal Statistical Society*. 172 (4). p. 749-769.

EVANS, M., PATTERSON, M. & O'MALLEY, L. (2001) The direct marketing-direct consumer gap: qualitative insights. *Qualitative Market Research: An International Journal*. 4(1). p.17-24.

EWING, T. (2008) Participation cycles and emergent cultures in an online community. *International Journal of Market Research*. 50(5). Web 2.0 Special Issue.

FABRIGAR, L. R., WEGENER, D. T., MACCALLUM, R. C., STRAHAN, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*. 4 (3). p .272–299. doi:10.1037/1082-989X.4.3.272.

FARAHAT, A., LOFARO, T., MILLER, J. C., RAE, G. & WARD, L. A. (2006) Authority rankings from HITS, PageRank, and SALSA: Existence, uniqueness, and effect of initialization. *SIAM Journal on Scientific Computing*. 27(4). p. 1181-1201.

- FAIRCHILD, A. J. & MACKINNON, D. P. (2009) A general model for testing mediation and moderation effects. *Prevention Science*. 10(2). p. 87-99.
- FARRAR, D. E. & GLAUBER, R. R. (1967) Multicollinearity in regression analysis: the problem revisited. *The Review of Economic and Statistics*. p. 92-107.
- FARQUHAR, J. & ROWLEY, J. (2006) Relationships and online consumer communities. *Business Process Management Journal*. 12 (2). p.162-177.
- FERRIN, D.L. & DIRKS, K.T. (2006) Direct and indirect effects of third-party relationships on interpersonal trust. *Journal of Applied Psychology*. 91(4). p. 870-883.
- FIELD, A. (ed.)(2009) *Discovering statistics using SPSS*. Sage publications.
- FIEDLER, M. & HARUVY, E. (2009) The lab versus the virtual lab and virtual field- An experimental investigation of trust games with communication. *Journal of Economic Behavior & Organization*.72(2). p. 716-724
- FINCH, B.J. (1999) Internet discussions as a source for consumer product customer involvement and quality information: an exploratory study.*Journal of Operations Management*.17. p.535-556.
- FINKBEINER, P., & DEAN, A. (2013) Capitalizing on social media to unleash the networked knowledge of a community of practitioners. *International Journal of Sales, Retailing and Marketing*. 1(2). p. 22-29.
- FLAVIAN, C., GUINALIU, M. & GURRA, R. (2006) The role played by perceived usability, satisfaction and consumer trust on website loyalty. *Information & Management*. 43(1). p. 1-14.
- FRITZ, M. S. & MACKINNON, D. P. (2007) Required sample size to detect the mediated effect. *Psychological science*. 18(3). p. 233-239.
- FROHLICH, N. (1974) Self-Interest or Altruism, What Difference? *The Journal of Conflict Resolution*. 18(1). p. 55-73.
- FU, F., LIU, L.H. & WANG, L. (2008) Empirical analysis of online social networks in the age of Web 2.0. *Physica A:Statistical Mechanics and its Applications*. 387(2). p. 675-684.
- FULK, J., HEINO, R., FLANAFIN, A.J., MONGE, P.R. & BAR, F. (2004) A test of the individual action model for organizational information commons. *Organization Science*.15 (5). p. 569-585.
- FULLER, J., MATZLER, K. & HOPPE, M. (2008) Brand community members as a source of innovation. *The Journal of Product Innovation Management*. 25(6). p. 608-619.
- GAGNE, M. (2009) A model of knowledge-sharing motivation. *Human Resource Management*. 48(4). p. 571.
- GALLAGHER, D., TING, L. & PALMER, A. (2008). A journey into the unknown; taking the fear out of structural equation modeling with AMOS for the first-time user. *The Marketing Review*. 8(3). p. 255-275.

- GANGULY, B., DASH, S.B. & CYR, D. (2009) Website characteristics, trust and purchase intention in online stores: - An empirical study in the Indian context. *Journal of Information Science and Technology*. 6 (2). p. 22-44.
- GANLE, D. & LAMPE, C. (2009) The ties that bind: Social network principles in online communities. *Decision Support Systems*. 47(3). p. 266-274.
- GARBARINO, E. & SLONIM, R. (2009) The robustness of trust and reciprocity across a heterogeneous U.S. population. *Journal of Economic Behavior & Organization*. 69(3). p. 226-240.
- GARCIA, H.F. (2007) Beware (and prepare for) the blogosphere era, *Strategy & Leadership*. 35(6). p. 51-53.
- GARLAND, R. (1991) The mid-point on a rating scale: Is it desirable. *Marketing bulletin*. 2(1). p. 66-70.
- GEFEN, D. & STRAUB, D. W. (2000) The relative importance of perceived ease of use in IS adoption: A study of e-commerce adoption. *Journal of the association for Information Systems*. 1(1). p. 8.
- GEFEN, D., STRAUB, D. & BOUDREAU, M. C. (2000) Structural equation modeling and regression: Guidelines for research practice. *Communications of the association for information systems*. 4(1). p. 7.
- GEORGE, D. & MALLEY, M. (2003) Using SPSS for Windows step by step: a simple guide and reference.
- GEORGE, D. & MALLEY, P. (2010) SPSS for Windows step by step. A simple study guide and reference. Boston, MA. USA: Ed.
- GERBING, D. W. & HAMITON, J. G. (1996) Viability of exploratory factor analysis as a precursor to confirmatory factor analysis. *Structural Equation Modeling: A Multidisciplinary Journal*. 3(1). p. 62-72.
- GHASEMI, A. & ZAHEDIASL, S. (2012) Normality tests for statistical analysis: a guide for non-statisticians. *International journal of endocrinology and metabolism*. 10(2). p. 486-489.
- GHOSHAL, G., & BARABASI, A. L. (2011). Ranking stability and super-stable nodes in complex networks. *Nature communications*. 2.p.394.
- GIBBS, G.R. (2002) *Qualitative data analysis: Explorations with Nvivo*. Buckingham, UK: Open University.
- GIFFIN, K. (1967) The contribution of studies of source credibility to a theory of interpersonal trust in the communication process. *Psychological bulletin*. 68(2). p. 104.
- GILSING, V.A. & DUYSTERS, G.M. (2008) Understanding novelty creation in exploration networks – Structural and relational embeddedness jointly considered. *Technovation*. 28(10). p. 693-708.
- GINTIS, H. (2007) A Framework for the Unification of the Behavioral Sciences. *Behavioral and Brain Sciences*. 30(01).p.1-61.
- GOFFIN, K. & KONERS, U. (2011) Tacit knowledge, lessons learnt, and new product development. *Journal of Product Innovation Management*. 28(2). p. 300-318.

GOLDSMITH, R.E. & FLYNN, L.R. (2005) Bricks, clicks, and pix: apparel buyers' use of stores, internet, and catalogs compared. *International Journal of Retail & Distribution Management*. 35 (4). p. 271-283.

GOLICIC, S. L. & DAVIS, D. F. (2012) Implementing mixed methods research in supply chain management. *International Journal of Physical Distribution & Logistics Management*. 42(8/9). p.726-741.

GORSUCH, R. L. (1990) Common factor analysis versus component analysis: Some well and little known facts. *Multivariate Behavioral Research*. 25(1). p. 33-39.

GRAHAM, J. M. (2006) Congeneric and (essentially) tau-equivalent estimates of score reliability what they are and how to use them. *Educational and Psychological Measurement*. 66(6). p. 930-944.

GRANDOCOLAS, U., RETTIE, R. & MARUSENKO, K. (2003) Web survey bias: sample or mode effect?. *Journal of marketing management*. 19(5-6). p. 541-561.

GRANOVETTER, M. S. (1973) The strength of weak ties. *American journal of sociology*. p. 1360-1380.

GRANOVETTER, M. (1978) Threshold models of collective behavior. *American journal of sociology*. p. 1420-1443.

GRANT, R. M. (1996) Toward a knowledge - based theory of the firm. *Strategic management journal*. 17(S2). p. 109-122.

GRAYSON, K., JOHNSON, D. & CHEN, D.F.R. (2008) Is firm trust essential in a trusted environment? How trust in the business context influences customers. *Journal of Marketing Research*. XLV. p. 241-256

GREENE, J. C. (2007) Mixed methods in social inquiry. (9). John Wiley & Sons.

GRIFFITH, D.A., KRAMPF, R.F. & PALMER, J.W. (2001) The role of interface in electronic commerce: consumer involvement with print versus on-line catalogues. *International Electronic Commerce*. 5(4). p.135-153.

GROMIE, J.G. & EWING, M.T. (2009) The rejection of brand hegemony. *Journal of Business research*. 62(2). p.218-230.

GRUEN, T.W., OSMONBEKOV, T. & CZAPLEWSKI, A.J. (2005) How e-communities extend the concept of exchange in marketing: An application of the motivation, opportunity, ability (MOA) theory. *Marketing Theory*. 5(1). p. 33-49.

GUILLEN, P., ENRIQUE, F. & PABLO, B.G. (2010) Inducing Efficient Conditional Cooperation Patterns in Public Goods Games: An Experimental Investigation. *Journal of Economic Psychology*. 31(6). p. 872-883.

GUPTA, P., YADAV, M.S.& VARADARAJAN, R. (2009) How task-facilitative interactive tolls foster buyers' trust in online retailers: A process view of trust development in the electronic Marketplace, *Journal of Retailing*. 85(2). p. 159-176.

GUPTA, S. & KIM, H.W. (2010) Value-driven Internet shopping: The mental accounting theory perspective. *Psychology & Marketing*. 27(1). p.13-35.

- GUTIERREZ, M.S.S., IZQUIERDO, C.C. & CABEZUDO, R.S.J. (2009) Product and channel-related risk and involvement in online contexts. *Electronic Commerce Research and Applications*.9(3). p. 263-273.
- GUY, R. F. & NORVELL, M. (1977) The neutral point on a Likert scale. *The Journal of Psychology*. 95(2). p. 199-204.
- HA, H.Y. (2004) Factors influencing consumer perceptions of brand trust online. *Journal of Product & Brand Management*. 13 (5). p.329-342.
- HAIR, J. F., BLACK, W. C., BABIN, B. J. & ANDERSON, R. E. (7th ed) (2010) *Multivariate Data Analysis: A Global Perspective*. London: Pearson Education, Inc.
- HAIR, J. F., RINGLE, C. M. & SARSTEDT, M. (2011) PLS-SEM: Indeed a silver bullet. *Journal of Marketing theory and Practice*. 19(2). p. 139-152.
- HAIR, J. F., GABRIEL, M. L. & PATEL, V. K. (2014). AMOS COVARIANCE-BASED STRUCTURAL EQUATION MODELING (CB-SEM): GUIDELINES ON ITS APPLICATION AS A MARKETING RESEARCH TOOL. *REMark: Revista Brasileira de Marketing*.13(2). p. 44-55.
- HAIR, J. F., SARSTEDT, M., HOPKINS, L. & KUPPELWIESER, G. V. (2014) Partial least squares structural equation modeling (PLS-SEM) An emerging tool in business research. *European Business Review*. 26(2). p. 106-121.
- HALLIDAY, J. (2011) Microsoft confirms \$8.5bn Skype deal. *The Guardian*. London. Retrieved 10 May 2011.
- HAMILTON, W. D. (1964) The genetical evolution of social behaviour. II. *Journal of theoretical biology*. 7(1). p. 17-52.
- HAN, Y.S., KIM, L. & CHA, J.W. (2009) Evaluation of user reputation on YouTube. *Online Communities and Social Computing*, 5621/2009.p.346-353.
- HANSEN, T., JENSEN, J.M. & SOLGAARD, H.S. (2004) Predicting online grocery buying intention: a comparison of the theory of reasoned action and the theory of planned behaviour. *International Journal of Information Management*. 24(6). p.539-550.
- HARRIS, L., & RAE, A. (2009) Social networks: the future of marketing for small business. *Journal of business strategy*.30(5). p.24-31.
- HARRIS, L.C., GOODE, M.M.H. (2004) The four levels of loyalty and the pivotal role of trust: a study of online service dynamics. *Journal of Retailing*. 80(2). p. 139-158.
- HASHIM, K., TAN, F. B. & ANDRADE, A. D. (2012) The Role of Trust and Commitment in Continuous Knowledge Contribution Intention within Business Online Communities. In *CONF-IRM 2012 Proceedings*.
- HARSANYI, J.C. (1978) Bayesian Decision Theory and Utilitarian Ethics. *The American Economic Review*. 68(2). p.223-228.

HASTIE, R., OSTROM, T. M., EBBSEBEN, E. B., WYER, R. S., HAMILTON, D. L., & CARLSTON, D. E. (2014) Person Memory (PLE: Memory): The Cognitive Basis of Social Perception (9). *Psychology Press*.

HAYTHORNTHWAITE, C. (2002). Strong, weak, and latent ties and the impact of new media. *The information society*. 18(5). p. 385-401.

HE, D.R., LIU, Z.H., & WANG, B.H. (ed.)(2008) *Complex Systems and Complex Networks*. Beijing: Higher Education Press.

HE, W. & WEI, K.K. (2009) What drives continued knowledge sharing? An investigation of knowledge-contribution and –seeking beliefs. *Decision Support Systems*. 46(4). p. 826-838.

HE, W., XU, L.D., MEANS, T. & WANG, P. (2009) Integrating Web 2.0 with the Case-based reasoning cycle: A systems approach. *Systems Research and Behavioural Science*. 26(6). p. 717-728.

HECHTER, M. & KANAZAWA, S. (1997) Sociological Rational Choice Theory. *Annual Review of Sociology*. 23. p. 191-214.

HEIDEMANN, J., KLIER, M. & PROBST, F. (2010) Identifying key users in online social networks: A pagerank based approach. Available from: Identifying key users in online social networks: A pagerank based approach. [Accessed :2010].

HERNANDEZ, B., JIMENEZ, J. & MARTIN, M. J. (2009) Consumer behaviour in electronic commerce: The moderating effect of e-purchasing experience. *Journal of Business Research*. 63(9). p. 964-971.

HES, D., & DU PLESSIS, C. (ed.) (2014) *Designing for Hope: pathways to regenerative sustainability*. Routledge.

HERTEL, G., KONRADT, U. & ORLIKOWSKI, B. (2004) Managing distance by interdependence: Goal setting, task interdependence, and team-based rewards in virtual teams. *European Journal of Work & Organizational psychology*. 13. p. 1–28.

HINZ, O. B. S., CHRISTIAN, B. & JAN, U. B. (2011) Seeding Strategies for Viral Marketing: An Empirical Comparison. *Journal of Marketing*. 75 (May). p. 55–71.

HOEHLE, H., HUFF, S. & GOODE, S. (2012) The role of continuous trust in information systems continuance. *Journal of Computer Information Systems*. 52(4). p. 1-9.

HOFFMAN, A. J., & BAZERMAN, M. H. (2007) Changing practice on sustainability: Understanding and overcoming the organizational and psychological barriers to action. *Organizations and the Sustainability Mosaic*. Crafting Long-Term Ecological and Societal Solutions. p. 84-105.

HÖGG, R., MECKEL, M., STANOEESKA-SLABEVA, K. & MARTIGONONI, R. (2006) Overview of business models for Web 2.0 communities. *Proceedings of GeNeMe*. p. 23-37

HOOPER, D., COUGHLAN, J., & MULLEN, M. (2008) Structural equation modelling: Guidelines for determining model fit. *Articles*. p. 2.

HOPWOOD, C. J. (2007) Moderation and mediation in structural equation modeling: Applications for early intervention research. *Journal of Early Intervention*. 29(3). p. 262-272.

HOSMER, L.T.(1995) Trust: The connecting link between organizational theory and philosophical ethics. *The Academy of Management Review*.20 (2). p.379-403.

HOSSAIN, L. & DE SILVA, A. (2009) Exploring user acceptance of technology using social networks. *Journal of High Technology Management Research*. 20(1). p.1-18.

HSU, C. L., & LIN, J. C. C. (2008) Acceptance of blog usage: The roles of technology acceptance, social influence and knowledge sharing motivation. *Information & Management*. 5(1). p. 65-74.

HSU, C. L., & LU, H. P. (2004) Why do people play on-line games? An extended TAM with social influences and flow experience. *Information & Management*.41(7). p. 853-868.

HSU, M.-H. & CHIU, C.-M. (2004) Predicting electronic service continuance with a decomposed theory of planned behaviour. *Behaviour & Information Technology*. 23(5). p.359 – 373.

HSU, M. H., JU, T. L., YEN, C. H., & CHANG, C. M. (2007). Knowledge sharing behavior in virtual communities: The relationship between trust, self-efficacy, and outcome expectations. *International journal of human-computer studies*. 65(2). p.153-169.

HSU, M.-H., YEN, C.-H., CHIU, C.-M. & CHANG, C.-M. (2006) A longitudinal investigation of continued online shopping behavior: An extension of the theory of planned behavior. *International Journal of Human-Computer Studies*. 64(9). p. 889-904.

HUANG, C.Y., CHOU, C.J. & LIN, P.C. (2010) Involvement theory in constructing bloggers' intention to purchase travel products. *Tourism Management*. 31(4).p.513-526.

HUFFAKER, D. (2010) Dimensions of Leadership and Social Influence in Online Communities. *Human Communication Research*. 36(4). p.593-617.

HUO, Q. & PALMER, A. (2015) Analysing Voluntary Contribution To Online Forums Using A Proposed Critical Mass Contribution Model. *Journal of Applied Business Research (JABR)*. 31(2). p. 687-700.

HURLEY, A. E., SCANDURA, T. A., SCHRIESHEIM, C. A., BRANNICK, M. T., SEERS, A., VANDENBERG, R. J. & WILLIAMS, L. J. (1997) Exploratory and confirmatory factor analysis: Guidelines, issues, and alternatives. *Journal of organizational behaviour*. 18(6). p. 667-683.

IACOBUCCI, D. (2009) Everything you always wanted to know about SEM (structural equations modeling) but were afraid to ask. *Journal of Consumer Psychology*. 19. p. 673-680.

IMHOF, M., VOLLMEYER, R. & BEIERLEIN, C. (2007) Computer use and the gender gap: The issue of access, use, motivation, and performance. *Computers in human behavior*, 23(6). p. 2823-2837.

INFOWOLD STAFF (2008) *Google buys Facebook Two Web powerhouses unite in \$25 billion deal*. Available from: <http://www.infoworld.com/d/adventures-in-it/google-buys-facebook-796>. [Accessed: 1th April 2008].

JANG, S., HONG, K., WOO BOCK, G. & KIM, I. (2002) Knowledge management and process innovation: the knowledge transformation path in Samsung SDI. *Journal of knowledge management*. 6(5). p. 479-485.

- JANSEN, B.J., SPINK, A. & SARACEVIC, T. (2000) Real life, real users, and real needs: a study and analysis of user queries on the web. *Information Processing & Management*. 36(2). p. 207-227.
- JEFFRIES, F. L., & BECKER, T. E. (2008). Trust, norms, and cooperation: Development and test of a simplified model. *Journal of Behavioral and Applied Management*. 9(3). p. 316.
- JIANG, J. J., KLEIN, G., & CARR, C. L. (2002) Measuring information system service quality: SERVQUAL from the other side. *MIS quarterly*. p. 145-166.
- JICK, T. D. (1979) Mixing qualitative and quantitative methods: Triangulation in action. *Administrative science quarterly*. 24(4). p. 602-611.
- JIN, X. L., CHEUNG, C. M., LEE, M. K., & CHEN, H. P. (2009) How to keep members using the information in a computer-supported social network. *Computers in Human Behavior*, 25(5), p. 1172-1181.
- JOHANNESSON, G., BJÖRBSSON, G. & GUDMUNDSSON, E. H. (2006) Afterglow light curves and broken power laws: a statistical study. *The Astrophysical Journal Letters*. 640(1). p. L5.
- JOHNSON, D. & GRAYSON, K. (2005) Cognitive and affective trust in service relationships. *Journal of Business research*. 58(4). p. 500-507.
- JONHNSON, D.S., BARDHI, F. & DUNN, D.T. (2008) Understanding how technology paradoxes affect customer satisfaction with Self-Service Technology: The role of performance ambiguity and trust in technology. *Psychology & Marketing*. 25(5). p. 416-443.
- KAHAI, S. S. & COOPER, R. B. (2003) Exploring the core concepts of media richness theory: The impact of cue multiplicity and feedback immediacy on decision quality. *Journal of Management Information Systems*. 20(1). p. 263-299.
- KAIKATI, A. M., & KAIKATI, J. G. (2004) Stealth marketing: how to reach consumers surreptitiously. *California Management Review*. 46(4). p. 6-22.
- KANT, I. (1948) *Groundwork of the Metaphysic of Morals*. N.Y: Harper Personnal.
- KANAZAWA, S. & SAVAGE, J. (2009) An evolutionary psychological perspective on social capital. *Journal of Economic Psychology*. 30(6). p.873-883.
- KANG, Y.S. & KIM, Y. J. (2006) Do visitors' interest level and perceived quantity of web page content matter in shaping the attitude toward a web site?. *Decision Support Systems*. 42(2). p. 1187-1202.
- KAPLAN, L.B., SZYBILLO, G.J. & JACOBY, J. (1974) Components of perceived risk in product purchase: A cross-validation. *Journal of Applied Psychology* .59(3). p. 287-291.
- KARAKAS, F. (2009) Welcome to World 2.0: the new digital ecosystem. *Journal of Business Strategy*. 23(4). P. 23-30.
- KASPRZAK, R. (2012) Diffusion in networks. *Journal of Telecommunications and Information Technology*. p. 99-106.
- KATZ, E. & LAZARSFELF, P. F. (1968) Personal influence. New York: Free Press.

KEELING, K., MCGOLDRICK, P. & BEATTY, S. (2009) Avatars as salespeople: Communication style, trust, and intentions. *Journal of Business Research*. 63(8). p. 793-800.

KELLER, E. & BERRY, J. (2006) Word-of-mouth: The real action is offline. *Advertising Age*. 77(49). p. 20.

KELTON, K., FLEISCHMANN, K.R. & WALLACE, W.A. (2008) Trust in digital information. *Journal of The American Society For Information Science and Technology*. 59 (3). p. 363-374.

KIESLER, S. & SPROULLI, L.(1992) Group decision making and communication technology. *Organization Behavior and Human Decision Processes*. 52(96). p. 96-123.

KIM, D.J., FERRIN, D.L. & RAO, H.R. (2008) A trust-based consumer decision-making model in electronic commerce: The role of trust, perceived risk, and their antecedents. *Decision Support Systems*. 44(2). p. 544-564.

KIM, J. O. & MULLER, C. W. (1978) Introduction to Factor Analysis: What It Is and How To Do It, Series: Quantitative Applications in the Social Sciences.

KIM, D.J., YUE, K.B., HALL, S.P. & GATES, T. (2009) Global diffusion of the Internet XV: Web 2.0 technologies, principles, and applications: A conceptual framework from technology push and demand pull perspective. *Communications of the Association for Information Systems*. 24(38). p.657-672.

KIM, H.S., PARK, J.Y. & JIN, B. (2008) Dimension of online community attributes: Examination of online communities hosted by companies in Korea. *International Journal of Retail & Distribution Management*. 36 (10). p. 812-830.

KIM, H.W. & GUPTA, S. (2009) A comparison of purchase decision calculus Between potential and repeat customers of an online store. *Decision Support Systems*. 47(4). p.477-487.

KIM, J.K., KIM, H.K., OJ, H.Y. & RYU, Y. U. (2009) A group recommendation system for online community. *International Journal of Information Management*. 30(3). p.212-219.

KIM, J. S., KAYE, J., & WRIGHT, L. K. (2001) Moderating and mediating effects in causal models. *Issues in Mental Health Nursing*. 22(1). p.63-75.

KIM, J.W., CHONNAM, J. C., QUALLS, W. & HAN, K. (2008) It takes a marketplace community to raise brand commitment: the role of online communities. *Journal of Marketing Management*.24(3-4). p.409-431.

KIM, J.Y. & SUNG, Y.J. (2008) Dimensions of purchase-decision involvement: Affective and cognitive involvement in product and brand. *Brand Management*. 26(8). p. 504-519.

KONANA, P. & BALASUBRAMANIAN, S. (2005) The Social-Economic-Psychological model of technology adoption and usage: an application to online investing. *Decision Support Systems*. 39(3). p.505-524.

KOUFARIS, M. & HAMPTON-SOSA, W. (2004) The development of initial trust in an online company by new customers. *Information & Management*. 41(3).p.377-397.

KUNEGIS, J., BLATTBER, M. & MOSER, C. (2013) Preferential attachment in online networks: Measurement and explanations. In Proceedings of the 5th Annual ACM Web Science Conference . ACM. May. p. 205-214

KOZINET, R.V. (1999) E-Tribalized Marketing: The strategic implications of virtual communities of consumption. *European Management Journal*.17(3). p. 252-264.

KOZINET, R.V. (2006) Click to connect: netnography and tribal advertising. *Journal of advertising research*. 46(3). p. 279-288.

KRACKHARDT, D., NOHRIA, N. & ECCLES, B. (1992) The strength of strong ties. Available from:

https://bdpems.wiwi.hu-berlin.de/portal/sites/default/files/Session%202_Krackhardt.pdf

KRANZ, S. (2010) Moral norms in a partly compliant society. *Games and Economic Behavior*. 68(1). p.255-274.

KRISHNAMURTHY, S. & KUCUK, S.U. (2009) Anti-branding on the internet. *Journal of Business Research*. 62(11). p.1119-1126.

KRUGMAN, H.E. (1965) The impact of television advertising: Learning without involvement. *Public Opinion Quarterly*. 29(3). p.349-356.

KUCUK, S. U. & KRISHNAMURTHY, S. (2007) An Analysis of consumer power on the Internet. *Technovation*. 27(1) .p.47-56.

KUMAR, R., NOVAK, J. & TOMKINS, A. (2010) Structure and evolution of online social networks. In KUMAR, R., NOVAK, J. & TOMKINS, A. (eds.) *Link mining: models, algorithms, and applications* (p. 337-357). New York : Springer.

KOH, J., KIM, Y. G., BUTLER, B. & BOCK, G. W. (2007) Encouraging participation in virtual communities. *Communications of the ACM*. 50(2). p. 68-73.

KOZINET, R. V. (2006) Netnography 2.0. In R. W. Belk (eds.) *Handbook of Qualitative Research Methods in Marketing* (p. 129-142). Cheltenham, UK: Edward Elgar.

KYRIACOU, A. P. (2011) Rational Irrationality and Group Size, The Effect of Biased Beliefs on Individual Contributions Towards Collective Goods. *American Journal of Economics and Sociology*. 70(1). p.109-130.

LAKSHMI, P.V. & MURUGAN, M.S. (2008) Socioeconomic characteristics and consumers' purchasing involvement. *ICFAI Journal of Consumer Behavior*.3(4). p. 69-79.

LAROCHE, M. (2009) Advances in internet consumer behavior and marketing strategy: Introduction to the special issue. *Journal of Business Research*. 63(9). p.1015-1017.

LARZELERE, R.E. & HUSTON, T.L. (1980) The dyadic trust scale: Toward understanding interpersonal trust in close relationships. *Journal of Marriage and the Family*. 42(3). p. 595-604.

LATORA, V. & MARCHIORI, M. (2007) A measure of centrality based on network efficiency. *New Journal of Physics*. 9(6). p. 188.

LAZARSELD, P. F., BERELSON, B. & GAUDET, H. (1968) The peoples choice: how the voter makes up his mind in a presidential campaign. Available from: <http://www.popline.org/node/517470>

LEARY, M. R. & FORSYTH, D. R. (1987). Attributions of responsibility for collective endeavors. *Review of Personality and Social Psychology*. 8. p. 167-188.

LEE N. & HOOLEY, G. (2005) The evolution of “classical mythology” within marketing measure development. *European Journal of Marketing*. 39(3/4). p. 365-385.

LEE, M.K.O. & TURBAN, E. (2001) A Trust Model for Consumer Internet Shopping. *International Journal of Electronic Commerce*.6(1). p. 75-91.

LEEFLANG, P.S.H., BIJMOLT, T.H.A., VAN DOORN J., HANSSENS, D.M., VAN HEERDE, H.J., VERHOEF, P.C. & WIERINGA, J.E. (2009) Creating lift versus building the base : Current trends in marketing dynamics. *International Journal of Research in Marketing*. 26(1). p.13-20.

LESKOVEC, J., HUTTENLOCHER, D., & KLEINBERG, J. (2010) Signed networks in social media. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* . p. 1361-1370. Available from: <http://arxiv.org/pdf/1003.2424.pdf>. [Accessed: 11th May 2010].

LESKOVEC, J., HUTTENLOCHER, D. P.& KLEINBERG, J. M. (2010, April). Governance in social media: A case study of the wikipedia promotion process. Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media. Available from: <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM10/paper/viewFile/1485/1841/>. [Accessed: 2010].

LEVY, M. (2009) Web 2.0 implications on knowledge management. *Journal of Knowledge Management*, 13(1). p. 120-134.

LEVIN, D. Z. & CROSS, R. (2004) The strength of weak ties you can trust: The mediating role of trust in effective knowledge transfer. *Management science*. 50(11). p. 1477-1490.

LEVIN, D. Z., CROSS, R., ABRAMS, L. C. & LESSER, E. L. (2002) Trust and knowledge sharing: A critical combination. *IBM Institute for Knowledge-Based Organizations*. p. 19.

LEWIS-BECK, M. S. (ed.)(1994) *Factor Analysis and Related Techniques*. London: Sage Publications.

LEWIS, T.G. (ed.)(2011). *网络科学理论与应用(Network Science :Theory and Applications)*. Beijing: China Machine Press.

LI, D. C. (2011) Online social network acceptance: a social perspective. *Internet Research*. 21(5). p. 562-580.

LI, L. (2004) Calculating the Confidence Intervals Using Bootstrap. *Presentation Notes, University of Toronto*. Available from : <http://math.usask.ca/~longhai/doc/talks/slide-bootstrap.pdf>. [Accessed: 12th November 2004].

LI, X., HESS, T.J. & VALACICH, J.S. (2008) Why do we trust new technology? A study of initial trust formation with organizational information systems. *Journal of Strategic Information Systems*.17(1). p.39-71.

LI, G., & YANG, X. (2014) Effects of social capital and community support on online community members' intention to create user-generated content. *Journal of Electronic Commerce Research* . 15(3). p.190.

LIBAI, B., MULLER, E., & PERES, R. (2013) Decomposing the value of word-of-mouth seeding programs: Acceleration versus expansion. *Journal of marketing research*. 50(2). p.161-176.

LIM, N. (2003) Consumers' perceived risk: sources versus consequences. *Electronic Commerce Research and Application*. 2(3). p.216-228.

LIN, H.F. (2007) The impact of Website quality dimensions on consumer satisfaction in the B2C E-commerce context. *Total Quality Management*. 18(4). p. 363-378.

LIN, H. F. (2007) Effects of extrinsic and intrinsic motivation on employee knowledge sharing intentions. *Journal of information science*. Available from:
<http://jis.sagepub.com/ttp://jis.sagepub.com/content/33/2/135>

LIN, H.F. (2008) Determinants of successful virtual communities: Contributions from system characteristics and social factors. *Information and Management*. 45(8). p.522-527.

LIN, N., COOK, C. & BURT, R.S. (2001) Social capital: Theory and research. New Brunswick, New Jersey :Transaction Publishers.

LING, K., BEENEN, G., LUDFORD, P., WANG, X., CHANG, K., LI, X. & KRAUT, R. (2005) Using social psychology to motivate contributions to online communities. *Journal of Computer-Mediated Communication*. 10(4). p. 00-00. Available from: <http://onlinelibrary.wiley.com/doi/10.1111/j.1083-6101.2005.tb00273.x/full>. [Accessed: 23rd June 2006].

LITTLE, T. D., CARD, N. A., BOVAIRD, J. A., PREACHER, K. J., & CRANDALL, C. S. (2007) Structural equation modeling of mediation and moderation with contextual factors. In LITTLE, T. D., CARD, N. A., BOVAIRD, J. A., PREACHER, K. J., & CRANDALL, C. S. (ed.) *Modeling contextual effects in longitudinal studies*. Routledge. Available from:

https://books.google.fr/books?id=jpRbBAAQBAJ&pg=PA207&dq=Structural+equation+modeling+of+mediation+and+moderation+with+contextual+factors&hl=fr&sa=X&ved=0CCEQ6AEwAGoVChMI1_Gc2oe4xwIVCtkaCh3iUQ9S#v=onepage&q=Structural%20equation%20modeling%20of%20mediation%20and%20moderation%20with%20contextual%20factors&f=false. [Accessed: 2012].

LITTLER, D. & MELANTHIOU, D. (2006) Consumer perceptions of risk and uncertainty and the implications for behaviour towards innovative retail services: The case of Internet Banking. *Journal of Retailing and Consumer Services*. 13(6). p.431-443.

LIU, L., LI, C., & ZHU, D. (2012) A new approach to testing nomological validity and its application to a second-order measurement model of trust. *Journal of the Association for Information Systems*. 13(12). p.950-975.

LO, S.K. & LIE, T. (2008) Selection of communication technologies –A perspective based on information richness theory and trust. *Technovation*. 28(3). p.146-153.

LOPEZ, E., PARSHANI, R., COHEN, R., CARMI, S., & HAVLIN, S. (2007) Limited path percolation in complex networks. *Physical review letters*. 99(18). p. 188701.

- LOPEZ-NICOLAS, C. & MOLINA-CASTILLO, F.J. (2008) Customer knowledge management and E-commerce: The role of customer perceived risk. *International Journal of Information Management*. 28(2). p.102-113.
- LOU, H., LUO, W. & STRONG, D. (2000) Perceived critical mass effect on groupware acceptance. *European Journal of Information Systems*. 9(2). p. 91-103.
- LU, J., YANG, J., & YU, C. S. (2013) Is social capital effective for online learning?. *Information & Management*. 50(7). p. 507-522.
- LU, Y.I., ZHAO, L. & WANG, B. (2009) From virtual community members to C2C e-commerce buyers: Trust in virtual communities and its effect on consumers' purchase intention. *Electron. Comm. Res. Appl.*. 9(4). p. 346-360.
- LUO, J.D. (2006) Particularistic trust and general trust: A network analysis in Chinese organizations. *Management and Organization Review*. 1(3). p. 437-458.
- LUO, X.M. (2002) Trust production and privacy concerns on the Internet: A frame work based on relationship marketing and social exchange theory. *Industry Marketing Management*. 31(2). p.111-118.
- LUTHI, L., PESTELACCI, E. & TOMASSINI, M. (2008) Cooperation and Community Structure in Social Networks. *Physica A : Statistical Mechanics and its Applications*. 387(4). p. 955-966.
- MAJEWSKI, M., & MALARZ, K. (2006). Square lattice site percolation thresholds for complex neighbourhoods. *arXiv preprint cond-mat/0609635*. Available from: <http://arxiv.org/abs/cond-mat/0609635>. [Accessed: 5th Febraury 2007].
- MADDEN, T.J, ELLEN, P.S. & AJZEN, I. (1992) A Comparison of the Theory of Planned Behavior and the Theory of Reasoned Action. *Personality and Social Psychology Bulletin*. 18(1). p. 3-9.
- MALHOTRA, N.K. (2007) *Marketing Research: An Applied Orientation*. NJ: Pearson Prentice-Hall.
- MALHOTRA N. K., AGARWAL, J. & PETERSON, M. (1996) Methodological issues in cross-cultural marketing research: A state-of-the-art review. *International marketing review*. 13(5). p. 7-43.
- MALLINCKRODT, A.W. T., WEI, M., & RUSSELL, D. W. (2006) Advances in testing the statistical significance of mediation effects. *Journal of Counseling Psychology*. 53(3). p. 372.
- MARTIN, G., REDDINGTON, M., KNEAFSEY, B.M. & SLOMAN, M. (2009) Scenarios and strategies for Web 2.0. *Education+Training*. 51(5/6). p.370-380.
- MARIN, S.S. & CAMARERO, C. (2008) Consumer trust to a Web site: Moderating effect of attitude toward online shopping. *CyberPsychology & Behavior*. 11(5). p.549-554.
- MARUYAMA, G. M. (ed.)(1997) *Basics of structural equation modeling*. USA:sage publications.
- MARWELL, G., OLIVER, P.E. & PRAHL, R. (1988) Social Networks and Collective Actions: A Theory of the Critical Mass. III. *American Journal of Sociology*. 94(3). p.502-534.
- MARWELL, G. & OLIVER, P. (1993) The critical mass in collective action. *Cambridge University Press*.

- MARZANO, G. & SCOTT, N. (2009) Power in destination branding. *Annals of Tourism Research*. 36 (2). p. 247-267.
- MASSEY, G. R., & DAWRS, P. L. (2007) The antecedents and consequence of functional and dysfunctional conflict between Marketing Managers and Sales Managers. *Industrial marketing management*. 36(8). p. 1118-1129.
- MAYER, A. (2009) Online social networks in economics. *Decision Support Systems*. 47(3).p. 169-184.
- MAZUREK, G. (2009) Web 2.0 implications on Marketing. *Management of Organizations: Systematic Research*. 51. p. 69-82.
- MCKNIGHT, D. H., CUMMINGS, L. L., & CHERVANY, N. L. (1998). Initial trust formation in new organizational relationships. *Academy of Management review* . 23(3). p. 473-490.
- MCKNIGHT, D. H., & CHERVANY, N. L. (2002) What trust means in e-commerce customer relationships: an interdisciplinary conceptual typology. *International journal of electronic commerce*. 6(2). p.35-60.
- MCKNIGHT, D.H., CHOUDHURY, V. & KACMAR, C. (2002) Developing and validating trust measure for e-commerce: An integrative typology. *Information Systems Research*. 13(3). p.334-359.
- MEINZEN-DICK, R. D.M, M. N. (2004) Methods for studying collective action in rural development. *Agric Sys*. 82(3). p.197–214.
- MELE, D. (2009) Integrating Personalism into Virtue-Based Business Ethics: The Personalist and the Common Good Principles. *Journal of Business Ethics*. 88(1).p. 227- 244.
- MESSINGER, P.R., STROULIA, E., LYONS, K., BONE, M., NIU, R.H., SMIRNOV, K. & PERELGUT, S. (2009) Virtual world-past, present, and future: New directions in social computing. *Decision Support Systems*. 47(3).p.204-208.
- MICHAELIDOU, N. & DIBB, S. (2008) Consumer involvement: a new perspective. *The Marketing Review*. 8(1). p. 83-99.
- MILLER, B. (2009) Community fundraising 2.0 – the future of fundraising in a networked society?. *International Journal of Nonprofit and Voluntary Sector Marketing*. 14(4). p. 365-370.
- MILLER, K.D., FABIAN, F. & LIN, S.J. (2009) Strategies for online communities. *Strategic Management Journal*. 30(3). p.305-322.
- MILES, M.B. & HUBERMAN, A.M. (ed.) (1994) *Qualitative data analysis: An expanded sourcebook*, Beverly Hills: Sage Publications.
- MINCER, M. & NIEWIADOMSKA-SZYNKIEWICZ, E. (2012) Application of social network analysis to the investigation of interpersonal connections. *Journal of Telecommunications and Information Technology*. p. 83-91.
- MINGERS, J. (2001) Combining IS research methods: towards a pluralist methodology. *Information systems research*. 12(3). p. 240-259.

MITCHELL, V.W. (1999) Consumer perceived risk: conceptualisations and models. *European Journal of Marketing*. 33 (1/2). p.163-195.

MITTAL, B. (1989) Measuring purchase-decision involvement. *Psychology & Marketing*. 6(2). p. 147-162.

MITTAL, B. (1998) Separating brand-choice involvement from product Involvement via consumer involvement profiles. *Advances in Consumer Research*. 15(1). p. 43-49.

MITZENMACHER, M. (2004) A brief history of generative models for power law and lognormal distributions. *Internet mathematics*. 1(2). p. 226-251.

MOHD KASSIM, N. & KADER MOHAMMED AHMED ABDULLA, A. (2006) The influence of attraction on internet banking: an extension to the trust-relationship commitment model. *International Journal of Bank Marketing*. 24(6). p. 424-442.

MOLLOY, M. & REED, B. (1995) A critical point for random graphs with a given degree sequence. *Random structures & algorithms*. 6(2 - 3). p. 161-180.

MOLONTAY, R. (2013) *Networks and fractals* (Doctoral dissertation, Department of Stochastics, Budapest University of Technology and Economics). Available from: <http://math.bme.hu/~molontay/Szakdolgozat.pdf>. [Accessed: 2013].

MONGKOLAJALA, M. H., PANICHPATHOM, S., & NGARMYARN, A. (2012) The Development of Tacit Knowledge Sharing Behaviour among Employees in Organizations. *International Journal of Business and Social Research*. 2(5). p. 158-163.

MOREIRA, A. A., ANDRADE, Jr, J. S., & AMARAL, L. A. N. (2002) Extremum statistics in scale-free network models. *Physical review letters*. 89(26). p. 268703.

MORGAN, R. M. & HUNT, S. D. (1994) The commitment-trust theory of relationship marketing. *the journal of marketing*. p.20-38.

MORREALE, S.P. & SHOCKLEY-ZALABAK, P.S. (2015) Organizational Trust in Cultures with a History of Distrust: A Qualitative Study of Polish and Russian Leaders' Perspectives and Experiences. *Journal of Intercultural Communication Research*. 44(1). p. 27-43.

MORTON, H. (2001) Computer-Mediated Communication in Australian Anthropology and Sociology. *Social Analysis Journal of Cultural and Social Practices*. 45(1). p. 3-11.

MOSSA, S., BARTHELEMY, M., STANLEY, H. E. & AMARAL, L. A. N. (2002) Truncation of power law behavior in "scale-free" network models due to information filtering. *Physical Review Letters*. 88(13).138701.

MORA, M. (Ed.). (2012). *Research Methodologies, Innovations and Philosophies in Software Systems Engineering and Information Systems*. IGI Global.

MUDAMBI, S.M., OLIVA, T.A. & THOMAS, E.F. (2009) Industrial marketing firms and knowledge transfer: Toward a basic typology of community structures. *Industrial Marketing Management*. 38(2). p. 181-190.

MULAIK, S. A., JAMES, L. R., VAN ALSTINE, J., BENNET, N., LIND, S. & STILWELL, C. D. (1989) Evaluation of goodness-of-fit indices for structural equation models. *Psychological bulletin*. 105(3). p. 430.

MULLER, D., JUDD, C. M. & YZERBYT, V. Y. (2005) When moderation is mediated and mediation is moderated. *Journal of personality and social psychology*. 89(6). p. 852.

MURPHY, P.E., LACZNIAK, G.R. & WOOD, G. (2007) An ethical basis for relationship marketing: a virtue ethics perspective. *European Journal of Marketing*. 41(1/2). p.37-57.

MYHILL, M., SHOEBRIDGE, M. & SNOOK, L. (2009) Virtual research environments-a Web 2.0 cookbook?. *Library Hi Tech*. 27(2). P. 228-238.

NAHAPIET, J. & GHOSHAL, S. (1998) Social capital, intellectual capital, and the organizational advantage. *The Academy of Management Review*. 23(2). p. 242-266.

NAHL, D. & TENOPIR, C. (1996) Affective and cognitive searching behavior of novice end-users of a full-text database. *Journal of the American Society for Information Science*. 47(4). p. 276-286.

NEUMANN, M., O'MURCHU, I., BRESLIN, J., DECKER, S., HOGAN, D. & MACDONAILL, C. (2004) Semantic social network portal for collaborative online communities. *Journal of European Industrial Training*. 29 (6). p.472-487.

NEWMAN, M. E. (2002). Assortative mixing in networks. *Physical review letters*. 89(20). p.208701.

NEWMAN, M. E. (2003). The structure and function of complex networks. *SIAM review*. 45(2). p.167-256.

NEWMAN, M. E. (2005) Power laws, Pareto distributions and Zipf's law. *Contemporary physics*. 46(5). p.323-351.

NEWMAN, M., BARABASI, A. L., & WATTS, D. J. (2006) *The structure and dynamics of networks*. USA:Princeton University Press. Available from: <http://press.princeton.edu/chapters/s8114.html>. [Accessed: 8th July 2007].

NEWMAN, M. E., STROGATZ, S. H., & WATTS, D. J. (2001) Random graphs with arbitrary degree distributions and their applications. *Physical review E*. 64(2). p. 026118.

NIELSEN, M.A. (2015) Neural Networks and Deep Learning. *Determination Press*.

Ng (2016) CS229 Machine learning course materials. Available from: <http://cs229.stanford.edu/materials.html>

NONAKA, I. (1994) A dynamic theory of organizational knowledge creation. *Organization science*. 5(1). p. 14-37.

NORRIS, M. & LECAVALIER, L. (2009) Evaluating the Use of Exploratory Factor Analysis in Developmental Disability Psychological Research. *Journal of Autism and Developmental Disorders*. 40 (1). p. 8–20. doi:10.1007/s10803-009-0816-2.

- NOWAK, G. J. & PHELPS, J. (1995) Direct Marketing and the use of individual-level consumer information: determining how and when “privacy” matters. *Journal of Direct Marketing*. 9(3). p.46-60.
- NOWAK, M.A., TARNITAN, G.E.& ANTAL,T. (2010) Evolutionary Dynamics in Structured Populations. *Philosophical Transaction of The Royal Society B:Biological Science*. 365(1537). p. 19-30.
- NOWAK, M.A. & ROGER, H. (ed.)(2011) *Super Cooperators- Evolution, Altruism and Human Behaviour or Why We Need Each Other to Succeed*. NY: Free Press.
- NUNNALLY, J. C., BERNSTEIN, I. H.& JOS MF, T.B.(1967). Psychometric theory, Vol 226. NewYork: McGraw-Hill
- OKLESHEN, C. & GROSSBART, S. (1998) Usenet groups, virtual community and consumer behaviour. *Advances in Consumer Research*. 25(1). p. 276-282.
- OKOLI, C. & OH, W. (2007) Investigating recognition-based performance in an open content community: A social capital perspective. *Information and Management*. 44(3). p. 240-252.
- O'LEARY-KELLY, S. W. & VOKURKA, R. J. (1998) The empirical assessment of construct validity. *Journal of operations management*. 16(4). p. 387-405.
- OLIVER, P.E. (1993) Formal model of collective action. *Annual Review of Sociology*. 19. p.271-300.
- OLIVER, P.E. & MARWELL, G. (1988) The paradox of group size in collective action: A theory of the critical mass.II. *American Sociological Review*. 53(1). p. 1-8.
- OLIVER, P.E. & MARWELLI, G. (2001) Whatever happened to critical mass theory? A retrospective and assessment. *Sociological Theory*. 19(3). p. 292-311.
- OLIVER, P., MARWELL, G. & TEIXEIRA, R. (1985) A theory of the critical mass. I. Interdependence, group heterogeneity, and the production of collective action. *American journal of Sociology*, 522-556.
- OLSON, M. (1965) The logic of collective action: Public goods and the theory of groups. Cambridge, Mass.
- O'REILLY, T. (2005) What is web 2.0. O'Reilly network.
- OSTROM, E. (ed.) (1990) *Governing the commons: The Evolution of Institutions for Collective Actions*. New York: Cambridge University Press.
- OSTROM, E. (2000) Collective action and the evolution of subjective norms. *The Journal of Economic Perspectives*. p.137-158.
- OSTROM, E. (2010) Analyzing Collective Action. *Agricultural Economics*. 41(1). p. 155-166.
- PALLANT, J. (ed. 3rd)(2007) *SPSS Survival Manual: A step by STEP guide to Data Analysis Using SPSS for Windows*. UK: Open University Press.
- PALMER, A & HUO, Q.Y. (2013) A study of trust over time within a social network mediated environment. *Journal of Marketing Management*. 29(15-16). p. 1816-1833.

PALMER, A., & KOENIG-LEWIS, N. (2009) An experiential, social network-based approach to direct marketing. *Direct Marketing: An International Journal*. 3(3). p. 162-176.

PARAMESWARAN, M. & WHINSTON, A.B. (2007) Research Issues in Social Computing. *Journal of the Association for Information Systems*. 8(6). p.336-350.

PARK, D.H. & KIM, S. (2008) The effects of consumer knowledge on message processing of electronic word-of-mouth via online consumer reviews. *Electronic Commerce Research and Applications*.7(4). p.399-410.

PARK, H. M. (2003) *Comparing group means: t-tests and one-way ANOVA using Stata, SAS, R, and SPSS*. USA: The University Information Technology Services (UITS) Centre for Statistical and Mathematical Computing, Indiana University. Available from:

http://stat.smmu.edu.cn/download/ebook/statistics_course.pdf. [Accessed : 2005].

PARMENTIER, G. & ROLLAND, S. (2009) Les consommateurs des mondes virtuels: construction identitaire et expérience de consommation dans Second Life. *Recherche et Application en Marketing*. 24(3). p. 113-129.

PASTOR-SATORRAS, R. & VESPIGNANI, A. (2001) Epidemic spreading in scale-free networks. *Physical review letters*. 86(14). p. 3200.

PATERSON, L. (2009) Online customer communities: Perspectives from customers and companies. *Business Information Review*. 26(1). p. 44-50.

PAVLOU, P.A. (2002) Institution-based trust in interorganizational exchange relationships: the role of online B2B marketplaces on trust formation. *Journal of Strategic Information System*.11(3). p. 215-243.

PAVLOU, P.A. & GEFEN, D. (2004) Building Effective Online Marketplaces with Institution-Based Trust. *Information Systems Research*.15(1). p.37-59.

PAYNE, A., STORBACKA, K., FROW, P. & KNOX, S. (2009) Co-creating brands: diagnosing and designing the relationship experience. *Journal of Business Research*. 62(3). p. 379-389.

PEARL, J. (ed.)(1998) *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Francisco: Morgan Kaufmann Publishers.

PEELEN, E., EKELMANS, C.F.W.&VIJN, P. (1989) Direct Marketing for establishing the relationships between buyers and sellers. *Journal of Direct Marketing*. 3(1). p.7-14.

PELISSIER, R. (2008) Business Research Made Easy. *Juta & Company Ltd*. Kenwyn, ZA.

PIRES, G., STANTON, J. & ECKFORD, A. (2004) Influences on the perceived risk of purchasing online. *Journal of Consumer Behaviour*.4(2). p.118-131.

PLANT, R. (2004) Online communities. *Technology in Society*. 26. p.51-65.

POSNER, E. (2000) *Law and Subjective norms*. Harvard University Press.

PRAHL, R., MARWELL, G. & OLIVER, P. E. (1991) Reach and selectivity as strategies of recruitment for collective action: A theory of the critical mass, V*. *Journal of Mathematical Sociology*.16(2). p.137-164.

PREACHER, K. J., & HAYES, A. F. (2008) Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior research methods*, 40(3). p. 879-891.

REACHER, K. J., RUCKER, D. D. & HAYES, A. F. (2007) Addressing moderated mediation hypotheses: Theory, methods, and prescriptions. *Multivariate behavioral research*. 42(1). p.185-227.

PREECE, J. (2001) Socialbility and usability: Twenty years of chatting online. *Behavior and Information Technology Journal*. 20(5). p.347-356.

PUNCH, K. F. (1998) Introduction to social research: Quantitative and qualitative approaches. Thousand Oaks. CA: Sage.

PURI, A. (2007) The Web of Insights: The Art and Practice of Webnography. *International Journal of Market Research*. 49(3). p.387-408.

QUINTAL, V.A., LEE, J. A. & SOUTAR, G.N. (2009) Risk, uncertainty and the theory of planned behaviour : A tourism example. *Tourism Management*. 31(6). p.797-805.

RABIN, M. (2001) Incorporating Fairness into Game Theory and Economics. *American Economic Review*. 83(5). p.1281-1302.

RAHMANDAD, H. & STERMAN, J. (2008) Heterogeneity and network structure in the dynamics of diffusion: Comparing agent-based and differential equation models. *Management Science*. 54(5). p. 998-1014.

RAPOPORT, A. (1957) Contribution to the theory of random and biased nets. *The bulletin of mathematical biophysics*. 19(4). p. 257-277.

RICHINS, M.L. & BLOCH, P.H. (1992) How enduring and situational involvement combine to create involvement responses. *Journal of Consumer Psychology*. 1(2). p. 143-153.

RICHINS, L.M. (1983) Word-of-Mouth by dissatisfied consumers: A pilot study. *Journal of Marketing*. 47(1). p. 68-78.

RIDINGS, C.M., GEFEN, D. & ARINZE, B. (2002) Some antecedents and effects of trust in virtual communities. *Journal of Strategic Information Systems*.11(3).p. 271-295.

RIDINGS, C. M., & GEFEN, D. (2004a) Virtual community attraction: Why people hang out online. *Journal of Computer-Mediated Communication*. 10(1). p.00-00. Available from:

<http://onlinelibrary.wiley.com/doi/10.1111/j.1083-6101.2004.tb00229.x/full>. [Accessed :23th June 2006].

RIDINGS, C. & WASCO, M. M. (2010) Online discussion group sustainability: Investigating the interplay between structural dynamics and social dynamics over time. *Journal of the Association for Information Systems*. 11(2). p. 95-121.

ROCHA, H.O. & GHOSHAL, S. (2006) Beyond Self-Interest Revisited. *Journal of Management Studies*. 43(3). p.585-619.

ROGERS, E. M. (ed.) (1995) Diffusion of innovation New York: Free Press.

ROUSSEAU, D. M., SITKIN, S. B, BURT, R. S. & CAMERER, C. (1998) Not so different after all: A cross-discipline view of trust. *The Academy of Management Review*.23. p. 393-404.

RYU, S., HO, S. H. & HAN, I. (2003) Knowledge sharing behavior of physicians in hospitals. *Expert Systems with applications*. 25(1). p. 113-122.

GRAHAM, R (ed.)(1997) *Game Theory-Introduction and Application*. Oxford: Oxford University Press.

SAHLI, A. B. & LEGOHEREL, P. (2015) The tourism Web acceptance model A study of intention to book tourism products online. *Journal of Vacation Marketing*. p.1356766715607589.

SALANCIK, G.R. & PFEFFER, J. (1978) A social information processing approach to job attitudes and task design. *Administrative Science Quarterly*. 23(2). p.224-253.

SAMUELSON, P.A (1954) The Pure Theory of Public Expenditure. *The Review of Economics and Statistics*. 36 (4). p. 387-389.

SANCHZE-FRANCO, M.J. & RONDAN-CATALUNA, F.J. (2010) Virtual travel communities and customer loyalty: Customer purchase involvement and web site design. *Electronic Commerce Research and Applications*. 9(2). p. 171-182.

SAYAMA, H., PESTOV, I., SCHMIDT, J., BUSH, B. J., WONG, C., YAMANOI, J., & GROSS, T. (2013) Modeling complex systems with adaptive networks. *Computers & Mathematics with Applications*. 65(10). p. 1645-1664.

SCHÖLKOPF, B. & SMOLA, A. J. (1998) From regularization operators to support vector kernels. *In Advances in Neural information processing systems*. 10. p. 343-349.

SCHOLSSER, A.E., WHITE, T.B. & LLOYD, S.M. (2006) Converting Web site visitors into buyers: how Web site investment increases consumer trusting beliefs and online purchase intentions. *Journal of Marketing*. 70(2). p. 133-148.

SCHUBERT, P. & GINSBURG, M. (2000) Virtual communities of Transaction: The Role of Personalization in Electronic Commerce. *Electronic Markets*.10 (1). p. 45-55.

SCHULTZ, C., SCHÖLER, L. & SKIERA, B. (2014) Not all fun and games: Viral marketing for utilitarian products. *Journal of Marketing*. 78(1). p. 1-19.

SCHUMACKER, RE. & LOMAX, R.G. (2004) A Beginner's Guide to Structural Equation Modelling , 2nd ed., London: Lawrence Erlbaum Associates.

SCHWARTZ, N., COHEN, R., BEN-AVRAHAM, D., BARABASI, A. L. & HAVLIN, S. (2002) Percolation in directed scale-free networks. *Physical Review E*. 66(1). p. 015104.

SEARLE, R. & BALL, K. (2004) The development of trust and distrust in a merger. *Journal of Managerial Psychology*. 19(7). p. 708-721.

SEO, W.J. & GREEN, B. C. (2007) The effect of Web cohesion, Web commitment, and attitude toward the Website on intentions to use NFL Teams' Websites. *Sport Management Review*. 10(3). p.231-252.

SHANG, R.A., CHEN, Y.C. & LIAO, H.J. (2006) The value of participation in virtual consumer communities on brand loyalty. *Internet Research*.16(4). p.398-418.

SHANKAR, V., URBAN, G.L. & SULTAN, F. (2002) Online trust: a stakeholder perspective, concepts, implications, and future directions. *Journal of Strategic Information Systems*.11(3). p. 325-344.

SHARRATT, M. & USORO, A. (2003) Understanding knowledge-sharing in online communities of practice. *Electronic Journal on Knowledge Management*. 1(2). p.187-196.

SHEEHAN, K.B. & HOY, M.G. (2002) Dimensions of privacy concern among online consumers. *Journal of Public Policy & Marketing*. 19(1). p. 62-73.

SHEN, H., WANG, L., XU, Q., LI, Y., & LIU, X. (2009) Toward a framework of innovation management in logistics firms: a systems perspective. *Systems Research and Behavioral Science*. 26(2). p.297-309.

SHEN, X. L., CHEUNG, C. M. & LEE, M. K. (2013) Perceived critical mass and collective intention in social media-supported small group communication. *International Journal of Information Management*, 33(5). p. 707-715.

SHETH, J. N., SETHIA, N. K.& SRINIVAS, S. (2011) Mindful consumption: a customer-centric approach to sustainability. *Journal of the Academy of Marketing Science*. 39(1). p. 21-39.

SHI, D.H.(ed.) (2011). 网络度分布理论 Theory of Network Degree Distributions. BEIJING: HIGHER EDUCATION PRESS.

SHMUELI, E., & ALTSHULER, Y. (2014) Temporal Dynamics of Scale-Free Networks. In *Social Computing, Behavioral-Cultural Modeling and Prediction* . p. 359-366. Springer International Publishing.

SHMUELI, G. (2010) To explain or to predict? *Statistical science*. p. 289-310.

SIEGEL, D. A. (2009) Social Networks and Collective Action. *American Journal of Political Science*. 53(1). p.122-138.

SLEDGIANOWSKI, D. & KULVIWAT, S. (2009) Using social network sites: The effects of playfulness, critical mass and trust in a hedonic context. *Journal of Computer Information Systems*. 49(4). p. 74-83.

SMALL, M., LI, Y., STEMLER, T. & JUDD, K. (2013) Super-star networks: Growing optimal scale-free networks via likelihood. arXiv preprint arXiv:1305.6429. Available from: <https://arxiv.org/abs/1305.6429>

SMITH, T. (2007) The social media revolution. *International Journal of Market Research*.51(4). p. 559-561.

SNIDAL, D. (1979) Public goods. *International Studies Quarterly*. 23(4). p. 532-566.

SNIEDER, R. & LARNER, K. (2009) The Art of Being a Scientist: A Guide for Graduate Students and their Mentors. *Cambridge University Press*.

SNIJDERS, C., MATZAT, U. & REIPS, U. D. (2012) Big data: Big gaps of knowledge in the field of internet science. *International Journal of Internet Science*. 7(1). p. 1-5.

SOBEL, M.E. (1982) Asymptotic confidence intervals for indirect effects in structural equation models. In: Leinhardt S, editor. *Sociological methodology*, 1985. Washington, DC: American Sociological Association. p. 290–312.

SOKOLOWSKI, R. (2000) Introduction to phenomenology. *Cambridge University Press*.

SONG, P.J., ZHANG, C., XU, Y.C. & HUANG, L.H. (2010) Brand extension of online technology products: Evidence from search engine to virtual communities and online news. *Decision Support Systems*.49(1). p. 91-99.

SOOMRO, T. R. & ASFANDYAR, D. (2010) Voice over Internet Protocol (VoIP): UAE Perspectives. *Asian Journal of Information Technology*. 9(3). p. 170-178.

SPAULDING, T.J. (2009) How can virtual communities create value for business. *Electronic Commerce Research and Applications*. 9(1). p. 39-48.

SPENCER, S. J., ZANNA, M. P. & FONG, G. T. (2005) Establishing a causal chain: why experiments are often more effective than mediational analyses in examining psychological processes. *Journal of personality and social psychology*.89(6). p. 845.

SRIDHAR, G. (2007) Consumer involvement in product choice: Role of perceived risk. *Decision*. 34(2). p.51-66.

STANDING, S., STANDING, C. & LOVE, P.E.D. (2010) A review of research on e-marketplaces 1997-2008. *Decision Support Systems*. 49(1). p. 41-51.

STECK, H. & JAAKKOLA, T. S. (2003) Bias-corrected bootstrap and model uncertainty. In *Advances in Neural Information Processing Systems* (p. None). Available from: <http://papers.nips.cc/paper/2427-bias-corrected-bootstrap-and-model-uncertainty.pdf>

STEENKAMP, J. B. E. & BAUMGARTNER, H. (2000) On the use of structural equation models for marketing modeling. *International Journal of Research in Marketing*. 17(2). p. 195-202.

STEVENS, J.R., CUSHMAN, F.A. & HAUSER, M.D. (2005) Evolving the Psychological Mechanisms for Cooperation. *Annual Review of Ecology, Evolution, and Systematics*. 36. p. 499-518.

STONE, M. (2009) Staying customer-focused and trusted: Web 2.0 and customer 2.0 in financial services. *Database Marketing & Customer Strategy Management*.16(2). p.101-131.

STRAUSS, A. & CORBIN, J. (ed.)(1998) Basics of qualitative research: Techniques and procedures for developing grounded theory. CA: Sage.

SUN, J.C. & YI, H.J. (2006) When brand attitudes affect the customer satisfaction-loyalty relation: The moderating role of product involvement. *Journal of Consumer Psychology*. 16(2). p.145-155.

SUN, K. (2004) *Explanation of log-normal distributions and power-law distributions in biology and social science*. Working Paper. Available from:

http://guava.physics.uiuc.edu/~nigel/courses/569/Essays_2004/files/sun.pdf. [Accessed: 6th May 2004].

SWAN, J. E., BOWERS, M.R., & RICHARDSON, L. D.(1999) Customer trust in the salesperson: An integrative review and meta-analysis of the empirical literature. *Journal of business research*. 44(2). p. 93-107.

SZULANSKI, G. (1996) Exploring internal stickiness: Impediments to the transfer of best practice within the firm. *Strategic management journal*. 17(S2). p. 27-43.

SZULANSKI, G. & JENSEN, R. J. (2006) Presumptive adaptation and the effectiveness of knowledge transfer. *Strategic Management Journal*, 27(10). p. 937-957.

TABACHNICK, B. G. & FIDELL, L. S. (ed.5th)(2007) Using Multivariate Statistics. London: Pearson Education, Inc.

TANAKA, J.S. (1987) “How big is enough?”:sample size and goodness of fit in structural equation models with latent variables. *Child development*. 58. p. 134 -146.

TANG, L. & LIU, H. (ed.)(2012). 社会计算 :社区发现和社会媒体挖掘(Community Detection and Mining in Social Media). Beijing : China Machine Press.

TAYLOR, C. & NOWAK, M. A. (2007) Transforming the dilemma. *Evolution*.61(10). p. 2281-2292.

TAYLOR, S., & TODD, P. A. (1995) Understanding information technology usage: a test of competing models. *Information systems research*. 6(2).p.144-176.

TEDDLIE, C. & TASHAKKORI, A. (2003) Major issues and controversies inthe use of mixed methods in the social and behavioral sciences. *Handbook of mixed methods in social & behavioral research*. p. 3-50.

TEDDLIE, C. & TASHAKKORI, A. (2009) Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences. *Sage Publications Inc*.

TELLIS, GERARD J., RAJESH, K. C.& DAVIS, S. A.(1999) In Search of Diversity: The Record of Major Marketing Journals. *Journal of Marketing Research*. 36(February). p.120–131.

TENENHAUS, M. (2008) Component-based structural equation modelling. *Total quality management*, 19(7-8). p. 871-886.

TSAI, C.H., ZHU, D.S., HO, B.C.T. & WU, D.D. (2010) The effect of reducing risk and improving personal motivation on the adoption of knowledge repository system. *Technological Forecasting & Social Change*. 77(6). p.840-856.

TSAI, H.T. & HUANG, H.C. (2007) Determinants of e-repurchase intention: An integrative model of quadruple retention drivers. *Information and Management*. 44(3).p.231-239.

TURRI, A. M., SMITH, K. H. & KEMP, E. (2013) Developing affective brand commitment through social media. *Journal of Electronic Commerce Research*. 14(3). p. 201.

VAN, B.P.J. & VAN, F. P.C. (2009) Instantiating global crisis networks: The case of SARS. *Decision Support Systems*.47(4). p.277-286.

VANCE, A., ELIE-DIT-COSAQUE, C. & STRAUB, D.W. (2008) Examining trust in information technology artefacts: The effects of system quality and culture. *Journal of Management Information Systems*. 24(4). p.73-100.

VANDERMARLIERE, B., KARAS, A., RYCKEBUSCGH, J. & SCHOORS, K. (2015) Beyond the power law: Uncovering stylized facts in interbank networks. *Physica A: Statistical Mechanics and its Applications*. 428. p.443-457.

VARGO, S. L. & LUSCH, R. F. (2009) Service-dominant logic: continuing the evolution. *Journal of the Academy of marketing Science*. 36(1).p .1-10.

VARLANDER, S. (2007) Online information quality in experiential consumption: An exploratory study. *Journal of Retailing and Consumer Services*. 14(5). p. 328-338.

VENKATESH, V., BROWN, S. A. & BALA, H. (2013) Bridging the qualitative-quantitative divide: Guidelines for conducting mixed methods research in information systems. *MIS quarterly*. 37(1). p.21-54.

VENKATESH, V., & GOYAL, S. (2010) Expectation disconfirmation and technology adoption: polynomial modeling and response surface analysis. *MIS quarterly*, p. 281-303.

VENKATRAMAN, N. (1989) The concept of fit in strategy research: toward verbal and statistical correspondence. *Academy of management review*. 14(3). p. 423-444.

VELOUTSOU, C. & MOUTINHO, L. (2009) Brand relationships through brand reputation and brand tribalism. *Journal of Business Research*. 62(3). p. 314-322.

VERBEKE, W. & BAGOZZI, R. P. (2000) Sales call anxiety: Exploring what it means when fear rules a sales encounter. *Journal of Marketing*.64(3). p. 88-101.

VON KROGH, G., NONAKA, I., & RECHSTEINER, L. (2012) Leadership in organizational knowledge creation: a review and framework. *Journal of Management Studies*. 49(1). p.240-277.

WALCZUCH, R. & LUNDGREN, H. (2004) Psychological antecedents of institution-based consumer trust in e-retailing. *Information & Management*. 42(1). p.159-177.

WANG, C.Y., YANG, H.Y. & CHOU, S.C. (2008) Using peer-to-peer technology for knowledge sharing in communities of practice. *Decision Support System*.45(3). p.528-540.

WANG, H.C., PALLISTER, J.G. & FOXALL, G.R. (2006) Innovativeness and Involvement as Determinants of Website Loyalty: III. Theoretical and managerial contributions. *Technovation*. 26(12). p. 1374-1383.

WANG, L. & DAI, G.Z. (ed.) (2009) 复杂网络的 Scale-free 性, Scale-free 现象及其控制. Beijing : Science Press.

WANG, X.F., Li, X., & CHENG, G.R. (ed.)(2006)复杂网络理论及其应用. Beijing: Qinghua University Press.

WANG, X.F., Li, X., & CENG, G.R.(ed.) (2012) 网络科学导论(Network Science: An Introduction). Beijing: Higher Education Press.

WANG,Y.F., LI, X.C., MAO, W.J., & WANG,T. (ed.)(2013)社会计算的基本方法与应用(Social computing: Methods and Applications). Zhe Jiang: ZHEJIANG UNIVERSITY RPRESS.

WASCO, M.M. & FARAJ, S. (2005)Why should I share? Examining social capital and knowledge contribution in electronic networks of practice. *MIS Quarterly*. 29(1). p.35-57.

WASCO, M.M., TEIGLAND, R. & FARAJ, S. (2009) The Provision of Online Public Goods: Examining Social Structure in an Electronic Network of Practice. *Decision Support Systems*. 47(3). p. 254-265.

WATTAL, S., RACHERLA, P., & MANDVIWALLAa, M. (2010) Network externalities and technology use: a quantitative analysis of intraorganizational blogs. *Journal of Management Information Systems*, 27(1).p.145-174.

WATTS, D. J. (2004) The" new" science of networks. *Annual review of sociology*. p. 243-270.

WATTS, D. J. & STROGATZ, S. H. (1998) Collective dynamics of ‘small-world’networks. *Nature*. 393(6684). p. 440-442.

WATTS, D.J. & Dodds, P.S. (2007) Influentials, Networks, and Public Opinion Formation. *Journal of Consumer Research*. 34 (4). p. 441–58.

WEBER, J.M, KOPELMAN, S. & MESSICK,D.M. (2004) A Conceptual Review of Decision Making in Social Dilemmas: Applying a Logic of Appropriateness. *Personality and Social Psychology Review*. 8(3). p. 281-307.

WEINBERGER, D. (2007) The real difference between the two 2.0s. *KM World*. 16(2). p.18 -20.

WEIJTERS, B., CABOOTER, E. & SCHILLEWAERT, N. (2010) The effect of rating scale format on response styles: The number of response categories and response category labels. *International Journal of Research in Marketing*. 27(3). p. 236-247.

WESTLAND, J.C. (2010) Critical Mass and Willingness to Pay for Social Networks. *Electronic Commerce Research and Applications*. 9(1). p. 6-19.

WESTLAND, J. C. (2010) Lower bounds on sample size in structural equation modeling. *Electronic Commerce Research and Applications*. 9(6). p. 476-487.

WESTON, R. & WESTON AND GORE, P. A. (2006) A brief guide to structural equation modeling. *The Counseling Psychologist*. 34(5). p.719-751.

WICKS, A.C. & BERMAN, S.L. (2004) The effects of context on trust in firm-stakeholder relationships: The institutional environment, trust creation, and firm performance. *Business Ethics Quarterly*. 14(1). p. 141-160.

WILLIAMSON, O.E. (1983) Credible Commitments: Using Hostages to Support Exchange. *The American Economic Review*. 73(4). p. 519-40.

WILSON, J. M., STRAUS, S.G. & MCEVILY, B. (2006) All in due time: The development of trust in computer-mediated and face-to-face teams. *Organisational Behavior and Human Decision Processes*. 99(1). p.16-33.

WRIGHT, S. (1922) Coefficients of inbreeding and relationship. *The American Naturalist*. 56(645). p. 330-338.

WRIGHT, K. B. (2005) Researching Internet - based populations: Advantages and disadvantages of online survey research, online questionnaire authoring software packages, and web survey services. *Journal of Computer - Mediated Communication*. 10(3). p. 00-00.

WRIGHT, R. T., CAMPBELL, D. E., THATCHER, J. B., & ROBERTS, N. (2012) Operationalizing multidimensional constructs in structural equation modeling: Recommendations for IS research. *Communications of the Association for Information Systems*. 30(1). p. 367-412.

WU, J.J., CHEN, Y.H. & CHUNG, Y.S. (2010) Trust factor influencing virtual community members: A study of transaction communities. *Journal of Business Research*. 63(9). p. 1025-1032.

WU, J.J. & TSANG, A.S. (2008) Factors affecting members' trust belief and behaviour intention in virtual communities. *Behaviour & Information Technology*. 27 (2). p. 115-125.

WU, W.P. & YANG, H.L. (2009) An empirical investigation of the relationship between moral intensity and ethical decision making in electronic commerce. *The International Journal of Organisation Innovation*. 2(2). p.195-210.

XIE, Z.(ed.) (2006) *Game Theories for Business and Economics*. Shanghai: University of Fudan Press.

XU, B., JONES, D.R. & SHAO, B. (2009) Volunteers' involvement in online community based software development. *Information & Management*. 46(3). p. 151-158.

YANG, S.Y. & VAN HECK, E. (2009) The Impact of multi-access technologies on consumer electronic auctions: A comparison of markets in China and the Netherlands. *European Management Journal*. 27(1). p.47-63.

YEN, H. R., HSU, S. H. Y. & HUANG, C. Y. (2011) Good soldiers on the Web: Understanding the drivers of participation in online communities of consumption. *International Journal of Electronic Commerce*. 15(4). p. 89-120.

YEOMANS, K. A. & GOLDER, P. A. (1982) The Guttman-Kaiser criterion as a predictor of the number of common factors. *The Statistician*. p. 221-229.

YI, Y.J., NATARAJAN, R. & GONG, T. (2011) Customer participation and citizenship behavioural influences on employee performance, satisfaction, commitment, and turnover intention. *Journal of Business Research*. 64(1). p. 87-95.

YIN, R.(ed.) (1984) *Case study research: Design & methods*. CA: Sage.

YOO, B. & DONTU, N. (2001) Developing a scale to measure the perceived quality of an Internet shopping site (SITEQUAL). *Quarterly Journal of Electronic Commerce*. 2(1). p. 31-47.

YUAN, K. H. & BENTLER, P. M. (1998) Normal theory based test statistics in structural equation modelling. *British Journal of Mathematical and Statistical Psychology*. 51(2). p. 289-309.

ZAICHKOWSKY, J.L. (1985) Measuring the involvement construct. *Journal of Consumer Research*.12. p. 341-352.

ZAİT, P. A., & BERTEA, P. S. P. E. (2011) Methods for testing discriminant validity. *Management & Marketing Journal*. 9(2).

ZHANG, P. (2011) 混合方法研究: 范式, 设计与质量标准. 天津大学学报: 社会科学版. 13(5). p.441-444.

ZHANG, J., ACKERMAN, M. S., & ADAMIC, L. (2007 May) Expertise networks in online communities: structure and algorithms. In *Proceedings of the 16th international conference on World Wide Web* . [Online] ACM. p. 221-230. Available from:

<http://www.wwwconference.org/www2007/papers/paper516.pdf>. [Accessed: 8-12 May 2007].

ZHANG, Y., FANG, Y. WEI, K. & CHEN, H. (2010) Exploring the Role of Psychological Safety in Promoting the Intention to Continue Sharing Knowledge in Virtual Communities. *International Journal of Information Management*. 30(5).p. 1-12.

Zhou, T. (2011) Understanding online community user participation: a social influence perspective. *Internet Research*. 21(1). p. 67-81.

ZHU, D., LI, X.B. & WU, S.L. (2009) Identity disclosure protection: A data reconstruction approach for privacy-preserving data mining. *Decision Support Systems*. 48(1). p.133-140.

ZIMMER, J. C., ARSAL, R. E., AL-MARZOUQ, M., & GROVER, V. (2010) Investigating online information disclosure: Effects of information relevance, trust and risk. *Information & management*. 47(2).p. 115-123.

ZUCKER, L.G. (1986) Production of trust: Institutional sources of economic structures, 1840-1920, *Research in Organizational Behavior*. 8. p. 63-111.

ZUO, Y.J. & PANDA, B. (2008) Two-level trust-based decision model for information in a virtual organization. *Decision Support Systems*. 45(2). p. 291-309.

ZVRIAN, M., GLEZER, C. & AVNI, I. (2006) User satisfaction from commercial web sites: The effect of design and use. *Information & Management*. 43(2). p. 157-178.